NTIS# PB 99-129371
HCFA CONTRACT # 500.93.0002
93-033
93-033-01-TJN 105

Abt Associates Inc.

55 Wheeler Street
Cambridge, Massachusetts
02138-1168

617 492-7100 *telephone*
617 492-5219 *facsimile*

Hampden Square, Suite 500
4800 Montgomery Lane
Bethesda, Maryland
20814-5341

301 913-0500 *telephone*
301 652-3618 *facsimile*

640 North LaSalle Street
Suite 400
Chicago, Illinois
60610-3781

312 867-4000 *telephone*
312 867-4200 *facsimile*

# Evaluation of Drug Use Review Demonstration Projects

## Contract No.: 500-93-0002

## Final Report

## October 28, 1998

*Submitted to:*
Jay P. Bae, PhD
Project Officer
Health Care Financing Administratio
Office of Research and Demonstratic
C-3-16-26 Central Building
7500 Security Boulevard
Baltimore, MD 21244-1850

*Submitted by:*
Abt Associates Inc.
55 Wheeler Street
Cambridge, MA 02138

# Table of Contents

1.0 **Background** . . . . . . . . . 1-1
  1.1 The Medicaid DUR Demonstrations . . . . . . . . . 1-2
    Project C.A.R.E. . . . . . . . . . 1-2
    Iowa OPDUR Demonstration Project . . . . . . . . . 1-3

  1.2 Approach to the External Evaluation . . . . . . . . . 1-3
    1.2.1 Components of the Evaluation . . . . . . . . . 1-4
    1.2.2 Evaluation Challenges . . . . . . . . . 1-5
    1.2.3 The Core Model . . . . . . . . . 1-6

  1.3 DUR and Medicaid Drug Policy: Trends Since OBRA 1990 . . . . . . . . . 1-7
    1.3.1 Medicaid Managed Care Penetration in the Four States . . . . . . . . . 1-7
    1.3.2 Medicaid Regulation of Prescription Drugs . . . . . . . . . 1-11

  1.4 Improving Prescribing . . . . . . . . . 1-11
    1.4.1 Introduction . . . . . . . . . 1-17
    1.4.2 The Role of Educational Interventions . . . . . . . . . 1-17
    1.4.3 Individual Educational Interventions . . . . . . . . . 1-18
    1.4.4 Group Educational Interventions . . . . . . . . . 1-19
    1.4.5 Feedback on Prescribing Practices . . . . . . . . . 1-20
    1.4.6 Clinical Pharmacist Systems . . . . . . . . . 1-22
    1.4.7 The Role of Incentives . . . . . . . . . 1-22
    1.4.8 The Role of Regulatory and Administrative Restrictions . . . . . . . . . 1-23
    1.4.9 Conclusions . . . . . . . . . 1-24

2.0 **Evaluation Research Tools** . . . . . . . . . 2-1

  2.1 The PSU Screener . . . . . . . . . 2-1
    2.1.1 History of the Penn State Screener . . . . . . . . . 2-1
    2.1.2 The PSU Screener: An Overview . . . . . . . . . 2-2
    2.1.3 Study Objective Served by the PSU Screener . . . . . . . . . 2-3
    2.1.4 Updating and Implementing the UM/PCPS Criteria . . . . . . . . . 2-3
    2.1.5 Porting and Verification of PSU Screener . . . . . . . . . 2-5
    2.1.6 Screen Definitions . . . . . . . . . 2-5

  2.2 The DUR Outcomes Database . . . . . . . . . 2-12
    2.2.1 Development of *The DUR Outcomes Bibliographic Database* . . . . . . . 2-12
    2.2.2 Selection of Candidate Clinical Outcomes . . . . . . . . . 2-14

3.0 **Drug Use and Prescribing Problems in Four State Medicaid Programs** . . . . . . . . . 3-1

  3.1 Background . . . . . . . . . 3-1
  3.2 Methods . . . . . . . . . 3-2
  3.3 Drug Use by Therapeutic Class . . . . . . . . . 3-4

## Table of Contents (continued)

3.4 DUR Screen Failure Results . . . . . . . . . . 3-5
  3.4.1 DUR Criteria Associated with Consistently Low Failure Rates . . . . . . . . . 3-5
  3.4.2 DUR Criteria Associated with Consistently High Failure Rates . . . . . . . . 3-6
  3.4.3 Comparative Findings by Drug Class . . . . . . . . . . 3-9
  3.4.4 Summary Measures of DUR Failure Rates . . . . . . . . . . 3-16

3.5 Summary and Conclusions . . . . . . . . . . 3-18

**4.0 Paying for Pharmacists' Cognitive Services: The Project C.A.R.E. Demonstration . . 4-1**

4.1 Background . . . . . . . . . . 4-1
4.2 Pharmacists and Prescribers: Attitudes Toward DUR and the Demonstration . . . . . 4-3
  4.2.1 Background . . . . . . . . . . 4-3
  4.2.2 OBRA 1990 Requirements . . . . . . . . . . 4-3
  4.2.3 Attitudes Toward Cognitive Services . . . . . . . . . . 4-4
  4.2.4 Payment for Provision of Cognitive Services . . . . . . . . . . 4-5
  4.2.5 The Pharmacist-Prescriber Relationship . . . . . . . . . . 4-6
  4.2.6 Pharmacist Attitudes Toward Work . . . . . . . . . . 4-7
  4.2.7 Recommendations for Improved Prescribing . . . . . . . . . . 4-7
  4.2.8 Conclusion . . . . . . . . . . 4-7

4.3 Project C.A.R.E.: Implementation . . . . . . . . . . 4-8
  4.3.1 Recruitment, Orientation and Training . . . . . . . . . . 4-8
  4.3.2 Sampling and Enrollment of Pharmacies and Cohort Maintenance . . . . . . 4-8
  4.3.3 Implementing the CS Payment Intervention . . . . . . . . . . 4-10
  4.3.4 Refinements and Modifications to the CS Payment Intervention . . . . . . . 4-10
  4.3.5 Organization, Staffing and Data Management . . . . . . . . . . 4-12
  4.3.6 Organizational and Staffing Issues: Problems and Solutions . . . . . . . . . 4-12
  4.3.7 Data Base Design and Implementation: Problems and Solutions . . . . . . . 4-12
  4.3.8 Data Flows . . . . . . . . . . 4-13
  4.3.9 Program Evaluation . . . . . . . . . . 4-14
  4.3.10 Conclusion . . . . . . . . . . 4-14

4.4 Findings on the Impact of the Cognitive Services Demonstration on Downstream Utilization and Costs . . . . . . . . . . 4-16
  4.4.1 Background . . . . . . . . . . 4-16
  4.4.2 Data and Methods . . . . . . . . . . 4-17
  4.4.3 Findings . . . . . . . . . . 4-25
  4.4.4 Discussion . . . . . . . . . . 4-28

4.5 Effects of Project C.A.R.E. on Prescription Use and Costs . . . . . . . . . . 4-30
  4.5.1 Data and Study Population . . . . . . . . . . 4-30
  4.5.2 Methods . . . . . . . . . . 4-31
  4.5.3 Findings . . . . . . . . . . 4-31
  4.5.4 Discussion . . . . . . . . . . 4-34

# Table of Contents (continued)

4.6 Effects of the Cognitive Services Demonstration on Screen Failures . . . . . . . . . . . 4-35
4.7 Outcomes Analyses — Methods and Data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4-48
4.8 The Impact of Project C.A.R.E. on Gastrointestinal Events . . . . . . . . . . . . . . . . . 4-53
4.9 The Impact of the Cognitive Services Demonstration on Adverse Clinical Outcomes Related to the Use of Antidepressants . . . . . . . . . . . . . . . . . . . . . . . . . 4-66
4.10 The Impact of the Cognitive Services Demonstration on Adverse Clinical Outcomes Related to the Use of Benzodiazepines . . . . . . . . . . . . . . . . . . . . . . . . 4-84
4.11 The Impact of Project C.A.R.E. on Adverse Outcomes Related to Inappropriate Prescribing of Cardiovascular Drugs . . . . . . . . . . . . . . . . . . . . . . 4-101
4.12 Costs and Cost-Effectiveness of Project C.A.R.E. . . . . . . . . . . . . . . . . . . . . . . . 4-121

**5.0 On-Line Prospective Drug Use Review: The OPDUR Demonstration** . . . . . . . . . . . 5-1

5.1 Background . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5-1
5.2 Iowa and Maryland Pharmacists Prescribers: Attitudes Toward DUR and the Demonstration . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5-2
5.3 The Iowa OPDUR Demonstration: Implementation . . . . . . . . . . . . . . . . . . . . . . 5-9

5.4 Findings on the Impact of the OPDUR Demonstration on Downstream Utilization and Costs . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5-12
    5.4.1 Background . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5-12
    5.4.2 Data and Methods . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5-12
    5.4.3 Findings . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5-21
    5.4.4 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5-28

5.5 Effects of the OPDUR Demonstration on Prescription Drug Use and Costs . . . . . 5-30
    5.5.1 Data and Study Population . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5-30
    5.5.2 Methods . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5-31
    5.5.3 Findings . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5-31
    5.5.4 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5-34

5.6 Effects of the OPDUR Demonstration Project on Screen Failures . . . . . . . . . . . . . 5-35
5.7 Outcomes Analyses — Methods and Data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5-49
5.8 The Impact of OPDUR on Gastrointestinal Events . . . . . . . . . . . . . . . . . . . . . . . 5-54
5.9 The Impact of the OPDUR Demonstration Project on Adverse Clinical Outcomes Related to the Use of Antidepressants . . . . . . . . . . . . . . . . . . . . . . . . . 5-67
5.10 The Impact of the OPDUR Demonstration Project on Adverse Outcomes Related to the Use of Benzodiazepines . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5-87
5.11 The Impact of the Iowa OPDUR Demonstration Project on Adverse Outcomes Related to Inappropriate Prescribing of Cardiovascular Drugs . . . . . . . . . . . . . . . 5-105
5.12 The Impact of the Maryland ProDUR Intervention for β-Agonist Inhaler Use on Elderly Dual-Eligible Medicaid Recipients . . . . . . . . . . . . . . . . . . . . . . 5-126
5.13 Costs and Cost-Effectiveness of the OPDUR Demonstration . . . . . . . . . . . . . . . 5-145

# Table of Contents (continued)

**6.0 Retrospective Drug Use Review** . . . . 6-1
- 6.1 Introduction . . . . 6-1
- 6.2 The Impact of RDUR — Iowa Medicaid DUR Board Letters on Misoprostol . . . . 6-3
- 6.3 The Impact of RDUR — Iowa Medicaid DUR Board Letters on Salmeterol . . . . 6-16
- 6.4 The Impact of RDUR — Maryland RDUR Letters on Zantac Use . . . . 6-30

**7.0 Conclusions** . . . . 7-1
- 7.1 Summary of Findings . . . . 7-1
- 7.2 Discussion . . . . 7-1

# Evaluation of Drug Use Review Demonstration Projects Final Report

## 1.0 Background

The Omnibus Budget Reconciliation Act of 1990 (OBRA 1990) required that all state Medicaid programs implement drug use review (DUR). This legislation also mandated demonstrations of innovative models of DUR, and in 1992, the Health Care Financing Administration (HCFA) funded two experiments managed by state Medicaid agencies. In Project C.A.R.E., Washington's Department of Social and Human Services collaborated with the University of Washington School of Pharmacy to test the effects of paying pharmacists to provide cognitive services (CS). The Iowa Medicaid OPDUR Demonstration Project, involving Iowa's Department of Human Resources, together with researchers from the University of Iowa, Drake University and the Iowa Pharmacists Association, tested on-line prospective drug use review (OPDUR).

In 1993, HCFA chose a team led by Abt Associates Inc. to conduct an external evaluation of both demonstrations, as well as targeted studies of the effectiveness of retrospective drug use review (RDUR).[1] In order to increase precision in quantitative analyses of the effects of OPDUR, the evaluation team recruited two additional states for the external evaluation: Maryland (designated a "co-experimental" state with an OPDUR system similar in structure to Iowa's), and Georgia (designated a "comparison" state with no OPDUR system in place). Because no other state was paying pharmacists for CS at the time the evaluation design was finalized, it proved to be impossible to find additional states to augment the evaluation of Washington's demonstration. Therefore, the external evaluation of CS payment effects relies only on Washington data.

This is the final report of the external evaluation.[2] The following sections of this chapter introduce the demonstrations and the evaluation design, review the Medicaid policy context within which the evaluation states operated, and review research evidence on programs designed to improve prescribing. Chapter Two describes two key research tools used in the evaluation: the "PSU screener" and a data base that relates prescribing problems codified in PSU screening criteria to adverse clinical outcomes, defined by ICD-9 diagnosis codes. Chapter Three presents baseline descriptive statistics on drug prescribing and prescribing problems in the four evaluation states. Chapter Four evaluates Project C.A.R.E. through a series of studies focused on CS payment effects in the Washington demonstration, including analyses of physicians' and pharmacists' attitudes and perceptions toward DUR, problems encountered in implementing the demonstration, and estimated effects of CS payment on prescription drug utilization, screen failures, "downstream" Medicare and Medicaid-reimbursed utilization, and clinical outcomes. In four separate outcomes analyses, the evaluation used epidemiological methods to study the effects of DUR on: 1) gastrointestinal bleeding associated with the misuse of NSAIDS, 2) a variety of outcomes associated with antidepressant prescribing problems (e.g. conduction disorders, falls), 3) hip fractures and other outcomes associated with benzodiazepine prescribing problems, 4) adverse outcomes associated with ACE inhibitor, calcium channel blocker and digoxin prescribing problems. This chapter concludes with a discussion of the costs and cost-effectiveness of cognitive services payment. Chapter Five repeats these analyses for an evaluation of OPDUR. This chapter also includes a fifth epidemiological study of the effects of Maryland's OPDUR program on outcomes associated with beta-agonist inhaler prescribing problems. Chapter Six reports on three studies of the effects of RDUR on prescription drug utilization. Two studies focus on Iowa interventions involving the drugs Misoprostol and Salmeterol, while one reports results from an H2 blocker (Zantac) intervention in Maryland. Chapter Seven concludes the report with a discussion of implications for Medicaid prescription drug policy.

---

[1] In addition to Abt Associates, the team included investigators from Pennsylvania State University, JEN Associates, Inc. and Casals and Associates.

[2] Both the Washington and Iowa demonstrations completed their own evaluations. See Christensen, D. B. et al., Payment of Pharmacists for Cognitive Services. Results of the Washington State C.A.R.E. Demonstration Project. HCFA Contract No. 11-C-90229. November 1996, and Chrischilles, E.A., The Iowa Medicaid OPDUR Demonstration Project. Annual Report. HCFA Cooperative Agreement No. 11-C-90232/7. September 1997.

## 1.1 The Medicaid DUR Demonstrations

Washington and Iowa designed the DUR demonstrations to improve drug prescribing by influencing pharmacists' behavior. However each demonstration tested a different assumption about what incentives are most effective in achieving the desired results.

### Project C.A.R.E.

Project C.A.R.E. paid pharmacists in a treatment group for providing cognitive services, defined as professional activities other than dispensing drugs that are related to dispensing and that are not routine, hypothesizing that payment would encourage more and more effective CS. To test the effects of paying for CS, Project C.A.R.E. recruited 200 pharmacies and randomly assigned geographic clusters of pharmacies to two groups: 110 pharmacies in group A clusters received payment for CS services and documentation, and 90 pharmacies in group B clusters received payment only for documenting CS provision. Clusters were created to capture patterns of prescriber use of pharmacies, minimizing "cross-over" of prescribers from one pharmacy type to another. Investigators also selected Group C, a "silent control" sample of pharmacies that received no CS-related payments and no notice that they had been selected. Project C.A.R.E. operated from February 1994 through September 1995, six months longer than originally planned.

To be eligible for payment under the demonstration, CS had to be related to dispensing a prescription drug to a Medicaid beneficiary, but not a "requisite part of dispensing" (i.e., accepting, interpreting, and clarifying a prescription order, preparing a prescription, delivering it to a patient), and had to conform to certain a priori assumptions about plausibility.[3] For each CS event, Project C.A.R.E. asked pharmacists to document the problem, the intervention, and the result. Investigators then determined what combinations of these three components "made sense." Experience gained during the demonstration expanded the set of codes determined to be reasonable. Finally, the demonstration limited payment to two CS per patient per day. Unlike Iowa OPDUR, Project C.A.R.E. did not focus pharmacists' attention on a specific list of drugs.

---

[3] Christensen, D.B., op. cit., page 17

For eligible cognitive services, Group A pharmacies received per service payments of $4.00 for a "brief encounter" (six minutes or less) or $6.00 for an "extended encounter" (over six minutes). Both A and B pharmacies received $40 per month for submitting CS documentation to the University of Washington. Group C pharmacies received no payments under the demonstration.

### Iowa OPDUR Demonstration Project

Iowa linked all pharmacies participating in the OPDUR demonstration to an on-line DUR screener that reviewed each Medicaid prescription drug claim as it was submitted electronically for payment. Pharmacies in the treatment group received messages that identified potential prescription problems. Control group pharmacies received no messages. Both treatment and control group pharmacies were asked to document all cognitive services provided, in order to describe demonstration effects on pharmacists behavior. Iowa did not pay pharmacists for providing or documenting cognitive services.

Unlike Washington, Iowa defined specific problems and drugs as the focus of its intervention. Iowa investigators worked with Unisys, Iowa's Medicaid fiscal agent, to customize ProDUR (TM), a software product of National Data Corporation. Operationally, the Iowa OPDUR screens were defined by criteria developed by the University of Maryland and the Philadelphia College of Pharmacy and Science. Criteria included drug-drug interactions, therapeutic duplications, and high daily dose for eight classes of drugs: angiotensin converting enzyme (ACE) inhibitors, calcium channel blockers, cardiac glycosides, benzodiazepines, antidepressants, antipsychotics, histamine H2 receptor antagonists, and nonsteroidal antiinflammatory agents (NSAIDS). Iowa also screened early refills for all drugs.

Iowa randomized pharmacy clusters to treatment and control groups. Of 223 pharmacies recruited for the demonstration, 110 were treatment and 113 control. The demonstration operated from June 1994 through June 1997.

## 1.2 Approach to the External Evaluation

HCFA asked the external evaluator to determine whether or not the demonstrations were effective 1) in improving drug use and 2) in reducing the negative consequences of drug misuse, measured in terms of adverse clinical events and health services utilization. Should one or both of the demonstrations show effectiveness, the evaluator was also expected to determine the costs of achieving these improvements, for comparison with other policies and programs designed to achieve these ends. This section describes the components of the evaluation, how the evaluator addressed methodological challenges, and the underlying model used to estimate effects.

### 1.2.1 Components of the Evaluation

There are three major components of the external evaluation. The basic evaluation methodology was replicated, separately, for Project C.A.R.E. and for the Iowa OPDUR Demonstration Project. These components include:

1) Descriptive contextual information. These sections included a discussion of problems encountered in implementing the demonstrations, based on reports from the demonstrations and site visits conducted by staff from the Abt Associates team; analyses of attitudes and perceptions of pharmacists and prescribers toward DUR and the demonstrations, based on a survey of pharmacists conducted in Washington in 1995 and in Iowa in 1996, as well as pharmacist and prescriber focus groups, held in all four states during 1995; baseline comparisons among the four evaluation states of Medicaid prescription drug utilization and problems.

2) Tests of aggregate demonstration effects. Three studies utilized a common statistical model to test for demonstration effects on

    a) *the volume and costs of prescription drug utilization* – In Iowa, where duplication and early refill are OPDUR explicit screens, but under-prescription is not, one would not expect to find the demonstration associated with increased drug utilization. However, in Washington, pharmacists are not guided in this manner by explicit screens, and so aggregate effects on drug utilization may be positive or negative.

    b) *the frequency of drug problems, measured as failures of PSU screener criteria* – The external evaluator used the PSU screener to measure effects on drug problems because it is the only reliable tool available for this purpose across all of the evaluation states. Screen failures in the eight PSU drug categories are expected to decline in Iowa, because Iowa screened against criteria defined for these drugs. Failures may have declined for these drugs in Washington, but there is no reason to believe that effects would be concentrated in them.

    c) *the volume and costs of "downstream" Medicare and Medicaid non-drug utilization* – Both interventions were expected to be effective in reducing the costs of downstream utilization, in particular, utilization associated with diagnoses identified with less-than-optimal prescribing (made operational by the PSU screen failure/outcome data base), emergency room visits, and physician office visits. However, there may be unintended effects on the volume of "optimal" prescribing, so that decreases in utilization of certain drugs result in increased downstream utilization in at-risk populations.

3) Focused tests of demonstration effects. The evaluation conducted five epidemiological analyses of the effects of the demonstrations on adverse clinical outcomes, in the context of specific prescribing problems. For example, by reducing the frequency of over-utilization of benzodiazepines, the demonstration may have effected a decrease in falls, an associated outcome. Because these studies focused on one clinical problem, modeling and choice of covariates was, to some extent, tailored to each problem. The study subjects were: NSAIDS/GI bleeding; benzodiazepines and falls; cardiac drugs and outcomes, beta-agonist inhalers and outcomes; and antidepressants and outcomes.

### 1.2.2 Evaluation Challenges

In every demonstration, challenges emerge when the demonstration design meets the realities of limited resources, institutional rigidities and random unforeseen events. These are discussed later, in Chapters Four (Project C.A.R.E.) and Chapter Five (Iowa OPDUR Demonstration Project). These demonstrations posed conceptual challenges as well.

***Defining subjects.*** Both Washington and Iowa randomly assigned pharmacy clusters to treatment and control groups, because the interventions were designed to affect pharmacists' behavior. The object of the demonstration, however, was to improve prescribing and reduce the frequency of preventable adverse outcomes for Medicaid beneficiaries. Therefore, the evaluators chose to analyze effects on beneficiaries rather than pharmacies or pharmacists. This might have posed a problem for the design, if beneficiaries filled prescriptions at both treatment and control pharmacies -- sampling in both demonstrations was structured to minimize crossover of physicians, with the expectation that beneficiary crossover would be minimized as well. This strategy was largely successful. In Washington, fewer than five percent of subjects used both Group A and B pharmacies during the study period; in Iowa, about seven percent of subjects used both Group A and B pharmacies during the study period. Because the proportion of crossovers was small in both demonstrations, these individuals were dropped from all analyses of demonstration effects.

Cluster sampling also introduced a potential source of variation that had to be controlled in the analyses in order to be sure that estimates of the demonstration effects were unbiased. In part, because the demonstrations sampled a fairly small number of clusters (37 in Washington and 151 in Iowa), the investigators were concerned that some cluster-level demographic or other factors might be correlated with treatment status, potentially producing estimates of demonstration effects with inaccurate standard errors. Measurable factors were controlled directly, through cluster-level averages that were included as regressors in multivariate models. In addition, investigators estimated the models using a version of the White/Huber variance estimation method that adjusts for correlation within clusters.

***Defining exposure.*** It is standard practice in evaluating social experiments to apply an "intent to treat" standard for defining exposure. That is, all members of the treatment group are exposed to the demonstration intervention as soon as the intervention is available, regardless of whether or not subjects actively participate in the demonstration. In the context of these demonstrations, one can define pharmacies as exposed by the date on which they were eligible for CS payment (Project C.A.R.E.) or ready to receive OPDUR messages (Iowa). Defining exposure for individual Medicaid recipients is less straightforward. Knowledgeable recipients could be aware of the timing and purpose of the demonstrations, without having any direct contact with a demonstration pharmacy, and this knowledge could affect their behavior. This conjecture seems implausible. More likely, recipients are exposed when first filling a prescription after the demonstrations have begun. This definition was adopted for the evaluation. Since pharmacies in both states entered the demonstrations in waves, recipient exposure reflects these recruitment patterns.

***Precision.*** Two factors suggest that demonstration effects, if they exist, will be small and difficult to find. First, the research literature implies that the frequency of prescribing errors with potentially harmful and costly consequences is considerably less than one percent [(Rupp et al. (1988, 1992)]. If, for example, the demonstration reduced prescribing error by 25 percent, from a frequency of 0.8 percent, the evaluator would be searching for an absolute effect of 0.2 percentage points. Second, the demonstrations both require a complex set of behavioral changes, the outcomes of which are uncertain. It is not sufficient for the subject pharmacist to respond to CS payment incentives or the information from an OPDUR message. The pharmacist must convince a prescriber, another pharmacist, or a recipient to act differently as well.

Appreciating these challenges, the external evaluator adopted two strategies to produce more precise estimates of demonstration effects. First, for the OPDUR demonstration, the evaluation recruited Maryland, which has an OPDUR program comparable to Iowa's, as a "co-experimental state." Georgia was recruited as a comparison state with no OPDUR program in place. For some analyses, it was believed that increased sample sizes would increase precision in the estimates of OPDUR effects. Second, the evaluation assessed demonstration effects on Medicaid sub-populations thought to be at particular risk of encountering prescribing problems. These included persons with diabetes, persons using multiple drugs (15+ in a reference month) and persons using more than three prescribers.

***Estimating effects.*** Several analyses, such as estimates of demonstration effects on multiple cardiac outcomes, produce multiple estimates of demonstration effects. If some of these estimates were found to be statistically significant and consistent with prior hypotheses regarding direction of effects, (but others were not), the investigator might be tempted to focus on the significant findings. In this case, however, the true Type I error (the chances of incorrectly concluding that there was an effect) is larger than its nominal value. For example, if the analysis included twenty independent hypothesis tests about the effects of the demonstrations on different cardiac outcomes, even if there were no true effects, at a five percent level of significance, one of twenty tests could be significant by chance. The evaluators agreed first to determine whether or not patterns of estimates posed this kind of problem. If a problem was found, accepted techniques for multiple tests would be employed. These include F-tests for combinations of all (or subsets) of estimates, or the Bonferroni adjustment for multiple tests.

### 1.2.3 The Core Model

Investigators applied a core model to estimate demonstration effects on several indicators, adjusted for other potentially confounding factors. The form this model took varied across individual studies in the evaluation. Later sections of this report describe nuances of differences embodied in each application. Every application, however, reflected the following conventions:

The dependent variable measures the effect of the demonstrations in a "follow-up" period defined to be when the demonstration was operational and when, for the treatment group, the subject was exposed to the demonstration through contact with a participating pharmacy. Exposure was captured by the treatment indicator, defined for most analyses at the recipient and pharmacy cluster level. The treatment indicator identified recipients that filled prescriptions at treatment pharmacies and not at control pharmacies during the demonstration period. The "effect" of the demonstration was estimated in the relationship between the treatment indicator and the dependent variable, adjusted for:

- **predisposing and need characteristics of the recipient,** limited to information captured from Medicare and Medicaid administrative records, which included age, gender, race, baseline medical conditions (diagnoses and procedures documented for baseline utilization), and

- **length of observation periods:** in order to adjust for the completeness of information in both the baseline and followup periods, investigators included measures for the length of both in the model. The baseline (pre-exposure) period varied from three to twelve months. The follow-up period varied from one to twenty months, driven both by the recipient's behavior (when did the recipient visit a treatment pharmacy) and by the dates when pharmacies entered the demonstration. The latter clearly affected observations both for treatment recipients and for control recipients, who were assigned to this group if they ever, during the demonstration, went to a control pharmacy and never went to a treatment pharmacy.

Social experiments never operate in a sterile, controlled environment. Among the important potential environmental effects on these demonstrations are state Medicaid policies. The next section summarizes changes in the four evaluation states in two key areas of Medicaid policy during the demonstrations: managed care enrollment and regulation of drug prescribing.

## 1.3 DUR and Medicaid Drug Policy: Trends Since OBRA 1990

Among the many Medicaid policies that could have affected the DUR demonstrations, managed care and drug regulation claim particular attention. At a practical level, utilization patterns of Medicaid managed care enrollees cannot be easily observed because, unlike fee-for-service, there is no standard system for submitting claims for payment (containing utilization information) to financial intermediaries or to HCFA.[4] Also, the system of capitation payment means that managed care organizations operate under different incentives from fee-for-service providers. Even if drug and other utilization data were readily available, managed care enrollment would be a major confounding variable in the analysis. The importance of capturing changes in prescription drug regulations should be obvious. Limits on refills, formularies, prior authorization and other policies can affect prescribing behavior, confounding measures of the demonstrations' effects.

The following section provides a brief overview of managed care penetration in the four states followed by a series of exhibits that describe state demographics and drug policies from 1990 to the present. The majority of the drug policy information was gathered from the National Pharmaceutical Council reference manual, and preliminary tables were sent to contacts in the four states for review.

### 1.3.1 Medicaid Managed Care Penetration in the Four States

The four evaluation states implemented managed care programs for Medicaid recipients in different ways. Exhibit 1.1 presents basic demographics and information on managed care penetration in the four evaluation states. Maryland citizens enjoy higher incomes (fewer families below the Federal Poverty Index than in the other three states). Both Maryland and Georgia have higher percentages of non-white individuals than Iowa and Washington. From information gathered through ongoing communication with contacts at the respective Medicaid bureaus, it was observed that as more Medicaid recipients enrolled in managed care, the pharmacy benefit was "carved-in"; i.e., prescription

---

[4] Even though, early in the enrollment of Medicaid recipients in managed care, pharmacy benefits were often carved out of the package of services covered under Medicaid capitated payments, effects of the demonstrations on other, non-drug utilization could not be observed.

Exhibit 1.1

HCFA DUR Demonstration Evaluation
Selected Demographic and Program Characteristics of the Demonstration States

| Attribute | Georgia | Maryland | Iowa | Washington |
|---|---|---|---|---|
| *POPULATION DATA, 1990* | | | | |
| Population | 6,478,216 | 4,781,468 | 2,776,755 | 4,866,892 |
| Race/Ethnicity* | | | | |
| • White | 71.0% | 71.0% | 96.6% | 88.5% |
| • African American | 27.0% | 24.9% | 1.7% | 3.1% |
| • American Indian | 0.2% | 0.3% | 0.3% | 1.7% |
| • Asian | 1.2% | 2.9% | 0.9% | 4.3% |
| • Hispanic (any race) | 1.7% | 2.6% | 1.2% | 4.4% |
| *INCOME* | | | | |
| Population below FPI (%state pop., 1989) | 14.2% | 8.1% | 11.1% | 10.6% |
| *MEDICAID MANAGED CARE, 1995* | | | | |
| Medicaid Recipients | 1,147,443 | 414,261 | 304,304 | 763,630 |
| Medicaid Managed Care Enrollment** | 139,748 | 320,106 | 262,790 | 458,179 |
| Recipients w/Pharmacy Benefits through MCO | All | All | None | All |

*Since Hispanics may be of any race, the sum of the percentages may exceed 100%.

**Represents Medicaid enrollees in health maintenance organizations, health insuring organizations, prepaid health plans, and primary care case management.

Sources: County and City Data Book 1994, National Pharmaceutical Council, Inc. 1997.

drugs were increasingly brought under the capitation payment rate. Most importantly, none of the four states made deliberate efforts to enroll elderly and disabled Medicaid recipients, the target populations for this evaluation, during the study period.

***Georgia.*** Georgia has brought Medicaid recipients into managed care in two ways: 1) through a primary care case management (PCCM) program, in which primary care providers, given a fee per case, deliver and coordinate all medical services; and 2) through a voluntary HMO program. Georgia initiated their PCCM program on a county-by-county basis and, as of the end of 1997, the program has been implemented statewide.

The PCCM program is mandatory for all Medicaid recipients except for those in nursing homes, other residential care, and recipients such as mothers with short-term Medicaid enrollment. Dual eligibles are enrolling on a voluntary basis. By January 1998, the program, called Georgia Better Health Care, had enrolled approximately 650,000 Medicaid recipients[5]. Pharmacy benefits remain under the auspices of the Department of Medical Assistance under the PCCM model.

The Department of Medical Assistance initiated its voluntary HMO program for AFDC and SSI populations in 1996. Unlike the PCCM program, this is a full capitation model and is inclusive of pharmacy benefits.

***Maryland.*** Between July and December 1997, Maryland enrolled Medicaid recipients in managed care under a Section 1115 waiver demonstration. SSI populations are now enrolled in managed care programs, but dual-eligibles are not. As for the pharmacy benefit, managed care programs are responsible for them with a few exceptions (i.e., anti retro-viral, specialty mental health prescriptions).

The waiver also establishes a special program called "The Rare and Expensive Case Management" (RECM) program for persons with such diseases as AIDS and hemophilia. These recipients are referred to a special RECM program as opposed to a managed care organization or HMO at the time of enrollment or diagnosis. Enrollment in this special program so far has reached approximately 850-900 participants.[6]

***Iowa.*** Iowa has several managed care initiatives underway, but the pharmacy benefit was excluded from these during the demonstration period in demonstration counties. Iowa began to offer an HMO option to Medicaid recipients in 1986 and initiated a PCCM system known as MediPASS in 1990. Recipient populations allowed to enroll in HMOs include most AFDC recipients (this program is known in Iowa as the Family Investment Program). As of 1998, SSI and SSI-related categories and dual eligibles had not been enrolled, although the state is considering including them in the future.[7] These populations do have carve-out benefits for mental health and substance abuse under managed care, but not for medical care.

***Washington.*** Healthy Options is a mandatory managed care program implemented beginning in 1993 for certain groups of Medicaid recipients which includes the pharmacy benefit. This program offers three managed care contracting models: (1) full risk-based capitation with a managed care company or HMO; (2) PCCM; and (3) a modified PCCM model through certified managed care organizations.

---

[5] Personal communication, Georgia Department of Managed Care.

[6] Personal communication, Maryland Department of Health and Hygiene.

[7] Personal communication, Iowa Department of Human Services.

The AFDC population was phased in during 1994. SSI and dual eligibles have not yet been moved into managed care.[8]

### 1.3.2 Medicaid Regulation of Prescription Drugs

With the exception of changes in formularies and prior authorization programs, and two initiatives in Washington described below, there was little drug policy activity in the four states during the study. Exhibits 1.2 through 1.5 describe the status from 1990-1997 of Medicaid policy variables that might have an impact on the evaluation of DUR programs in the four study states, including maximum days supply (Exhibit 1.2), maximum number of prescriptions (Exhibit 1.3), and maximum number of refills (Exhibit 1.4). As noted (Exhibit 1.5), Iowa had in place a prior authorization program that expanded during the demonstration, a trend that will make it difficult to detect DUR effects, especially for H2-receptor antagonists, NSAIDS and benzodiazepines.

Washington experienced short-lived policy changes that disrupted early phases of the demonstration. In February 1994, the State implemented a supplemental rebate program that effectively restricted drugs sold by companies that refused to sign supplemental rebate agreements with the State. Pharmacists had to telephone an informational line to determine whether or not a prescribed drug was restricted. Negative reaction to the program and to Medicaid was substantial, resulting in repeal of the policy in July 1995. Also, in January 1994, the State mandated a $1 copayment for most Medicaid prescriptions. There was considerable confusion among pharmacists and, combined with the rebate program, more animosity toward the Medicaid program. This policy was repealed in April 1994.

## 1.4 Improving Prescribing

DUR provides tools and incentives for improved drug use. The Medicaid DUR demonstrations were designed to achieve this through the actions of pharmacists, directed toward prescribers, other pharmacists and recipients. Encouraged by financial incentives or better information, pharmacists were expected to intervene more frequently to resolve prescribing or compliance problems. Most OPDUR screening criteria (with the possible exception of early refills) target prescription problems (drug-drug interaction, therapeutic duplication) rather than patient compliance. Project C.A.R.E. investigators reported that pharmacists directed over two-thirds of all CS interventions to prescribers and patients, and were equally likely to intervene with these two groups.

Arguably, the prescriber is key to any successful DUR intervention, even when the immediate problem seems to stem from patient non-compliance. Therefore this section of Chapter 1 will review research evidence on the effectiveness of programs designed to change prescriber behavior.

---

[8] Personal communication, Washington State Department of Social and Health Services.

**Exhibit 1.2**

**HCFA DUR Demonstration Evaluation**
**State Medicaid Policy on Maximum Allowable Day's Supply for Covered Prescriptions**

| Year | Georgia | Maryland | Iowa | Washington |
|---|---|---|---|---|
| 1990 | Physicians are encouraged to prescribe a 30 day supply, and Medicaid will pay for a maximum of a 31 day supply. | The amount per prescription at one time is limited to a less than 34-day supply except for specific maintenance drugs for chronic conditions, where up to a 100-day supply may be dispensed at one time. | Maintenance drugs may be supplied in 90-day quantities. All other drugs are limited to a 30-day supply | Minimum of 30 days for maintenance drugs (drugs for chronic conditions). After that authorization is needed. |
| 1991 | | | | |
| 1992 | | | | |
| 1993 | | | | |
| 1994 | | | | |
| 1995 | | | | |
| 1996 | | | The practitioner shall prescribe a 30-day supply, unless therapeutically contraindicated. Exception: Maintenance drugs | Maximum of 34 days for maintenance drugs. |
| 1997 | | | | |

Source: National Pharmaceutical Council, Inc. (1990-1997).
Pharmaceutical benefits under state Medicaid programs. Reston, VA: Author.

**Exhibit 1.3**

## HCFA DUR Demonstration Evaluation
## State Medicaid Policy on Maximum Allowable Number of Prescriptions

| Year | Georgia | Maryland | Iowa | Washington |
|---|---|---|---|---|
| 1990 | Six prescriptions per month per recipient except by prior authorization. The anxiolytic Benzodiazepines allowed 3 prescriptions per calender year without prior authorization. | No restrictions listed. | No restrictions listed. | No restrictions listed. |
| 1991 | | | | |
| 1992 | Six prescriptions per month for children, five prescriptions per month for adults, except by prior authorization. | | | |
| 1993 | | | | |
| 1994 | | | | |
| 1995 | | | | |
| 1996 | H2's and other antiulcer drugs are restricted to 2 prescriptions per lifetime at full dose therapy. Non-sedating (anti)histamines are limited to one prescription a year. Toradol and Dolobid are also limited to 1 prescription per year without prior authorization. | | | |
| 1997 | | | | |

Source: National Pharmaceutical Council, Inc. (1990-1997).
Pharmaceutical benefits under state Medicaid programs. Reston, VA: Author.

**Exhibit 1.4**

### HCFA DUR Demonstration Evaluation
### State Medicaid Policy on Maximum Number of Refills

| Year | Georgia | Maryland | Iowa | Washington |
|---|---|---|---|---|
| 1990 | Refills good for 2 years on legends, not controlled. Refills of medication dispensed within the same month of service are reimbursed at only the drug cost, with the following exceptions: antibiotics, C-II-IV's, Clozaril, and oral contraceptives. A second dispensing fee is only paid for the exceptions noted. | The maximum authorized number of refills is two. Refills may not be dispensed after 100 days from the date of the original prescription. The original prescription and its refills may not exceed a 100-day supply. | No restrictions in policy therefore state law would indicate 11 refills for non-controlled medications or 5 refills for controlled medications. | No more than 2 refills in any 30-day period unless prescription and refills are in amount of 100s. |
| 1991 | | | | |
| 1992 | | | | No more than 2 refills in 30-day period. Refills only in amounts of 100 |
| 1993 | | | | No more than one refill in 30-day period. Refills only in amounts of 100. |
| 1994 | | | | No more than 2 refills in 30-day period. Refills only in amounts of 100. |
| 1995 | | | | |
| 1996 | | | No restrictions listed. | No more than 2 refills in 30-day period except for antibiotics, anti-asthmatics, schedule II and III drugs, antineoplastics, topicals, and any propoxyphene. |
| 1997 | | | | |

Source: National Pharmaceutical Council, Inc. (1990-1997). Pharmaceutical benefits under state Medicaid programs. Reston, VA: Author.

**Exhibit 1.5**

**HCFA DUR Demonstration Evaluation**
**Restricted Drug Formularies in State Medicaid Programs**

| Year | Georgia | Maryland | Iowa | Washington |
|---|---|---|---|---|
| 1990 | The Controlled Medical Assistance Drug List. | Open formulary. | No formulary. Most OTC drugs, amphetamines, laxatives, and prescription multiple vitamins require prior authorization. General exclusions include: products used for anorexia or weight gain; drugs used for cosmetic purposes or hair growth; drugs used to promote smoking cessation; DESI drugs; and drugs used for fertility purposes. | Restricted formulary includes approx. 2,900 listing by drug product name, quantity, dosage form and strength. Excluded products include OTC drugs. |
| 1991 | | Open formulary with certain exceptions such as products for which Federal Financial Participation is not allowed, i.e., products whose manufacturers have not signed rebate agreements. | | |
| 1992 | | | No formulary. Prior authorization required for: clozapine, H2-receptor antagonists, sulcrafate, omeprazole, misoprostal (beyond 90 days), single source NSAIDS and benzodiazepines. | No formulary. Specifications for "non-formulary" drugs are now listed for "non-contract" drugs because the formulary policy was discontinued. |
| 1993 | Allowable OBRA '90 exclusions | | | |

**Exhibit 1.5 (continued)**

**HCFA DUR Demonstration Evaluation**
**Restricted Drug Formularies in State Medicaid Programs**

| Year | Georgia | Maryland | Iowa | Washington |
|---|---|---|---|---|
| 1994 | | | No formulary. Most non-legend drugs not provided. Prior authorization required for: growth hormones, topical acne products, non-sedating antihistamines, cephaloxin hydrochloride monohydrate, epoetin, filgrastim, and legend multiple vitamins | |
| 1996 | | | | Open formulary. General exclusions include: agents for anorexia or weight gain; fertility; cosmetic or hair growth; smoking cessation; DESI drugs and vitamins. Prior authorization is required for growth hormones and brand drugs scheduled by the DEA department of narcotics and tranquilizers. |
| 1997 | Open formulary. General exclusions include all OBRA '90 exclusions and agents to promote hair growth, fertility drugs, and anorexiants. Ambien limited to 18 tabs/month, Imitrex tabs limited to 18 tabs/month, Imitrex injections limited to 9 injections/month. Stadol nasal spray 54m/month or 18 spray units. | | Open formulary. General exclusions include: products used for anorexia or weight gain; drugs used for cosmetic purposes or hair growth; drugs used to promote smoking cessation; DESI drugs; drugs used for fertility purposes. | Open formulary. General exclusions include: agents for anorexia or weight gain, fertility, cosmetic or hair growth, smoking cessation, DESI drugs, and vitamins. |

Source: National Pharmaceutical Council, Inc. (1990-1997). Pharmaceutical benefits under state Medicaid programs. Reston, VA: Author.

### 1.4.1 Introduction

It is well-known that drug therapy problems can lead to costly adverse health outcomes. In a study of board and care facilities in ten states, Spore et. al. (1997) find that between 20-25 percent of elderly residents have at least one inappropriate prescription; other research has indicated that drug misadventures account for up to 30 percent of all hospital admissions. Johnson and Bootman (1995) estimate that the yearly health system cost of drug-related morbidity and mortality is $76.6 billion. Most recently, Phillips et.al. (1998) report that the proportion of outpatient deaths attributable to medication errors has increased dramatically, from one in 539 outpatient deaths in 1983 to one in 131 outpatient deaths in 1993. This suggests that increased monitoring of drug use is needed, particularly in the outpatient setting where prescribers have less control over medication administration.

Several programs and policies designed to optimize and regulate physicians' prescribing practices have been implemented. The objective of this section is to survey recent studies that evaluate the effectiveness of these programs and policies. The majority of the literature in this area focuses on educational interventions designed to change prescriber behavior. Correspondingly, most studies reviewed here are evaluations of a variety of educational interventions. A few researchers also have addressed the consequences of regulations and incentives meant to improve prescribing behavior. This research is summarized in two separate sections following the discussion of educational interventions.

Throughout the literature review, the study design of each evaluation usually is noted. The results of randomized trials and interrupted time series studies are discussed in greater detail compared to pre/post and other observational studies. Because there is little information on the impact of these policies and programs on long-term outcomes, the focus of the discussion is short-run prescribing improvements, costs, and health outcomes.

This review is divided into three sections based on the intervention method under evaluation: (1) educational interventions (including individual, group, feedback and clinical pharmacy programs); (2) incentive-based interventions; and (3) administrative restrictions or regulations. These categories are somewhat artificial because in some cases, the intervention methods overlap (i.e. educational intervention with regulations). As will become clear, the research indicates that a single, effective method of modifying prescriber behavior does not exist. Instead, different clinical settings, prescribers and goals seem to require different interventions.

### 1.4.2 The Role of Educational Interventions

In general, continuing education programs aimed at physicians are diverse, widespread and fairly effective. Davis et.al. (1995), in a review of continuing medical education program evaluations, find that 70 percent of the studies reviewed report some positive (albeit often quite small) effects on physician behavior. The continuing medical education programs reviewed included short courses, academic detailing, reminder systems and audits with feedback. [Davis et.al. (1995)] Importantly, these researchers only reviewed evaluations that were randomized controlled trials or otherwise scientifically rigorous. [Davis et.al. (1995)]

Many researchers have focused on the impact of educational interventions specifically designed to alter physicians' prescribing behavior. Although there is evidence that a range of factors determine inappropriate prescribing,[9] the rationale for educational interventions is that providers need more information in order to select safe, cost-effective drugs for their patients. In this literature review, evaluations of educational programs are grouped into four broad categories: (1) individual educational interventions; (2) group educational interventions; (3) feedback on prescribing practices; and (4) clinical pharmacy services.[10] These categories include a range of interventions that vary widely in intensity, effectiveness and cost. In some cases, there is considerable evidence that a specific form of the intervention is very effective (i.e., face-to-face prescription counseling from another physician). Other types of educational interventions have not been evaluated rigorously or offer mixed evidence of success.

### 1.4.3 Individual Educational Interventions

Some forms of individual educational interventions have proved very successful in improving prescriber behavior. In particular, several rigorous studies demonstrate the effectiveness of academic detailing. Academic detailing is defined as face to face, informational discussions between trained counselors and prescribers. Although the objective of the academic detailing visit is to educate the prescriber about specific drug problems (e.g., an adverse interaction between two drugs), the discussion also can uncover important non-clinical factors in drug prescription, such as the influence of patient demands and advertising. The academic detailing approach is based in part on the highly effective sales approach used by most pharmaceutical companies worldwide. [Soumerai and Avorn (1990)] [11]

In a widely cited, randomized controlled study, Avorn and Soumerai (1983) evaluate the effectiveness of two educational interventions: (1) a set of two academic detailing meetings with highly trained clinical pharmacists; and (2) mailing of a series of informational, printed materials. The objective of the educational interventions was to reduce the usage of two ineffective drugs and one potentially toxic drug [Avorn and Soumerai (1983)]. Nine months after the beginning of the intervention, providers who had participated in academic detailing had reduced usage of the targeted drugs by 14 percent compared to a control group.[12] The mailing intervention appeared to have no significant effect on prescriber behavior [Avorn and Soumerai (1983)].

Schaffner et al. (1983), in another controlled trial, also find that academic detailing is effective and mailed materials are ineffective in changing prescriber behavior. The objective of this intervention was to reduce prescriptions of antibiotics that had potentially adverse consequences for young children.

---

[9] Schwartz et. al. (1989), in a study of physicians' prescribing behavior in four states, find that physicians attribute their own inappropriate prescribing decisions not to a lack of information but rather to patient demand, clinical experience that contradicts scientific findings, and the use of particular drugs as placebos [Schwartz et. al. (1989)].

[10] Soumerai et al. (1989) categorize interventions in a somewhat more detailed fashion, including 1) dissemination of printed educational materials; 2) reporting of patient-specific listings of prescribed medications; 3) group education; 4) feedback of prescribing practices; 5) reminders at the time of prescribing; 6) face-to-face outreach; and 7) clinical pharmacy systems (i.e. in-house pharmacist who participates in prescribing decisions).

[11] Soumerai and Avorn (1990) cite a 1984 Wall Street Journal report stating that at that time pharmaceutical companies were spending at least $5000 per physician per year on detailing efforts in the United States.

[12] During the observation period, the control group reduced use of targeted drugs by 9 percent while the intervention group reduced use by 23 percent [Avorn and Soumerai (1983)].

Unlike Avorn and Soumerai (1983), Schaffner et al. (1983) sent either a pharmacist or a physician to prescribers' offices for the detailing visit. Interestingly, only the prescribers who received physician detailing visits changed their behavior, reducing usage of the targeted drugs by 54 percent compared to the control group in some cases [Schaffner et al. (1983)]. In a later follow-up study, these researchers found that the reduction in targeted drug usage lasted for two years after the intervention [Ray et al. (1985)]. In two other controlled trials, Peterson and Sugden (1995) and Ray et al. (1986) report that individual, academic detailing significantly improves physicians' prescriptions of allopurinol and diazepam.

In a recent study, Lin et al. (1997) use quasi-experimental methods to evaluate a year-long effort to improve physicians' treatment of major depression. The program included individual academic detailing as well as other educational activities. Although treatment practices did improve initially, after six months prescribing behavior had reverted back to baseline behavior [Lin et al. (1997)]. Similarly, in a controlled trial, Ray et al. (1987) find that a single, face-to-face academic detailing visit had no discernable effect on physicians' use of anti-psychotics in nursing homes [Ray et al. (1987)].

It is likely that the success of individual academic detailing varies by certain program characteristics such as the academic background of the detailer, the length of program, and the setting of the intervention.[Schaffner et. al. (1983), Soumerai and Avorn (1990)] Nevertheless, based on the success of initial academic detailing efforts, many researchers and policy makers have embraced this method of modifying prescriber behavior. [Lipton and Bird (1993); Greco and Eisenberg (1993)]

Educating individual physicians is often expensive, however, and the cost-effectiveness of academic detailing has not been established. In a later article, Soumerai and Avorn (1986) address the issue of cost-effectiveness of the academic detailing program described in an earlier article by the same authors [(Avorn and Soumerai (1983)]. These authors estimate the costs of the academic detailing effort using process data collection and assumptions about salaries, benefits and travel time. They find that at approximately $94.00 per physician (in 1980 dollars), their program would yield net benefits between $66,000-$111,000 per thousand physicians if the program were expanded to serve 10,000 physicians [Soumerai and Avorn (1986)].

Despite this evidence of considerable cost savings, there is very little if any other scientific information about the cost-effectiveness of individual educational interventions such as academic detailing. As Soumerai and Avorn (1990) suggest, it may be most cost-effective to target physicians with particularly inappropriate prescribing habits [Soumerai and Avorn (1990)]. For other physicians, less costly methods of education may be more cost-effective.

### 1.4.4 Group Educational Interventions

Although studies of individual academic detailing programs are more widely cited, many researchers have assessed the effectiveness of group-based educational interventions designed to change prescriber behavior. Group educational interventions can range from low-cost, informational mailings to relatively expensive group detailing programs. These methods potentially offer the effectiveness of individual academic detailing at a lower cost.

In a controlled trial, Avorn et al. (1992) evaluate the effectiveness of a pharmacist-led, educational intervention designed to reduce use of psychoactive drugs by nursing home patients. The program consisted of individual informational sessions aimed at targeted physicians as well as group educational sessions for nurses and nurse assistants [Avorn et al. (1992)]. The program proved to be very effective, reducing psychoactive drug prescriptions by 32 percent in the experimental group compared to a reduction of 14 percent in the control group [Avorn et al. (1992)]. Furthermore, Avorn et al. (1992) report that experimental group patients experienced improvements in memory and other clinical outcomes [Avorn et al. (1992)].

Similarly, Diwan et al. (1995) find that in a randomized control trial, a four-session pharmacist-led group detailing program significantly improved the prescribing practices of primary care physicians. In this study, appropriate prescriptions per month for lipid lowering drugs increased by 20 percent in the intervention group compared to a slight decrease in the control group [Diwan et al. (1995)]. Guitierrez et al. (1994) also find that a group educational intervention is effective in improving prescriber behavior. In this controlled trial conducted in Mexico, physicians participated in five classes designed to promote oral rehydration therapy use and other appropriate care for pediatric diarrhea. The researchers report significant differences between the experimental and control groups that persisted 18 months after the intervention [Guitierrez et al. (1994)].

Compared to interventions designed to improve patient outcomes, there is less evidence that interventions specifically targeted at lowering costs are effective. Schroeder et al. (1984), for example, find that a group education program aimed at reducing costs was ineffective. In this controlled trial, fourth-year medical students and first-year residents participated in four lectures and chart audits. The program was led by a variety of professionals, including physicians. Despite an increase in the intensity of the program in its second year, the intervention and control groups displayed no significant differences in mean charges. [Schroeder et al. (1984)]. Appelgate et al. (1983) also evaluate a group educational intervention designed to lower costs. Although these authors do find significant effects, the pre/post study design they used has limitations.

Less intensive group educational efforts (i.e., mailings, short courses) also show more limited evidence of effectiveness compared to intensive group education efforts and individual academic detailing. Although some studies without control groups suggest that the dissemination of printed educational materials to groups can modify prescriber behavior, controlled trials have shown that educational mailings generally are ineffective when used alone [Soumerai and Avorn (1986), Soumerai et al. (1989), Schauberger et al. (1996), Lundborg (1997), Lipton and Bird (1993), Schaffner et. al. (1983)]. Similarly, clinical practice guidelines designed to change the behavior of groups of physicians are quite ineffective unless they are combined with other educational methods [Greco and Eisenberg (1993)]. Among group educational interventions, therefore, the more intensive, face-to-face programs demonstrate the most potential for success.

### 1.4.5 Feedback on Prescribing Practices

The educational interventions described thus far focus on the modification of general prescription practices. Educational interventions can also be designed to offer information at the time of prescription or feedback on past prescribing practices, often implemented as part of retrospective drug use review (RDUR) interventions. Because this type of educational intervention focuses on the prescriber's specific behavior, it may be more effective than a general education program.

Frazier et al. (1991), for example, find that prescription feedback reports combined with weekly reminders and the distribution of a drug price manual were effective in changing prescriber behavior. In this controlled trial, the intervention group (medical interns) prescribed less expensive drugs compared to the control group, but it is not clear which aspect of the program was most effective. [Frazier (1991), 116] Britton and Lurvey (1991) also find that feedback on prescribing practices is helpful. In this controlled trial, clinical pharmacists reviewed charts and made prescription recommendations for a group of patients using at least five drugs per month. This system of medical profile review led to lower medication usage and lower medication costs in the intervention group [Britton and Lurvey (1991)]. It is not clear, however, what effect this intervention had on patient outcomes.

There also is some evidence that Medicaid drug utilization review intervention letters and feedback letters in the HMO setting improve prescribing practices. These targeted letters usually inform physicians of their inappropriate use of a particular drug with a particular patient or set of patients.

Many RDUR studies have focused on interventions designed to improve the prescribing of anti-ulcer drugs. Guo et al. (1995) find that in Alabama, intervention group physicians who received a DUR letter achieved lower anti-ulcer drug reimbursement rates compared to a control group [Guo et al. (1995)]. This study did not address patient outcomes. In Texas, Rascati et al. (1996) also focus on the use of Medicaid DUR letters to modify anti-ulcer medication prescribing. Although these researchers also find that the letters significantly reduced duplicative anti-ulcer treatment in the intervention group, no information on patient outcomes is presented [Rascati et al. (1996)]. Similarly, in a quasi-experimental study in Wisconsin, Zimmerman et al. (1994) find that an intervention group of physicians who received DUR letters reduced their use of histamine-2 receptor antagonists (H2RA). Although the researchers found that hospitalization rates for gastrointestinal problems were lower in the experimental group compared to the control group, these hospitalization rates were very low for both groups (0 to 1.5 percent) [Zimmerman et al. (1994)]. Nevertheless, it is noteworthy that Zimmerman et. al. (1994) assessed patient outcomes.

In a later quasi-experimental study of a Wisconsin Medicaid letter intervention intended to reduce inappropriate use of dipyridamole, Collins et. al. (1997) again find the intervention is effective. In this case, the authors are able to compare the effects of letters sent to both pharmacists and physicians to the effects of letter sent to physicians only. They find that letters sent to both pharmacists and physicians are more effective than letters sent to physicians only [Collins (1997)].

Most recently, Brufsky et. al. (1998), using an interrupted time series with comparison method, study the impact of a multi-faceted feedback effort to shift anti-ulcer drug therapy from ranitidine to cimetidine. The intervention, which took place in an HMO setting, involved educational memos as well as feedback on prescribing of anti-ulcer drugs. The objective of the intervention was to reduce the costs and maintain the quality of ulcer drug therapy. The authors estimate that the intervention reduced total H2RA costs by $1.06 million in the first year without affecting hospitalization rates for gastrointestinal problems [Brufsky et al. (1998)]. This evidence is consistent with Schectman et. al. (1995), who also find effects of feedback letters on H2 Blocker prescribing in the HMO setting.

At this time, computer based prescriber feedback systems have not been rigorously evaluated [Soumerai and Lipton (1995)]. Vedsted et al. (1997) and Keys et al. (1995) offer some preliminary, conflicting evidence about the efficacy of specific computer systems. As Soumerai and Lipton (1995) suggest, it is important to conduct more research in this area to ensure that the spread of these systems benefits patients [Soumerai and Lipton (1995)]. Overall, then, researchers have found that feedback interventions are effective in modifying prescribers' behavior. At this time, however, there is little evidence that prescription feedback improves patients' outcomes or reduces costs at the hospital or clinic level.

### 1.4.6 Clinical Pharmacist Systems

It is possible that prescriber education is best accomplished through frequent informal and formal interactions with pharmacists. Ideally, clinical pharmacist systems incorporate the pharmacist as part of a prescribing team that selects the most appropriate drug therapy for the patient. Anecdotal information and pre/post studies support the idea that clinical pharmacist systems can effectively modify prescriber behavior [Hawkey et al. (1990), Peterson and Goldberg (1989), Herfindal et al. (1985), Chrischilles et al. (1984), Sulick and Pathak (1996), Brown et al. (1979) and Lobeck et al. (1989)].

Controlled trials in this area offer some mixed evidence that clinical pharmacy services may reduce inappropriate prescriptions and lower drug costs. Stergachis et al. (1987) find that when clinical pharmacists interacted formally (i.e. lectures) and informally (i.e. consultations) with physicians, prescription volumes of targeted drugs changed only slightly and costs were not affected [Stergachis et al. (1987)]. Chrischilles et al. (1989), on the other hand, find that physician/patient appointments that took place in a clinical pharmacist-staffed clinic resulted in more appropriate prescriptions compared to appointments in a clinic without a pharmacist [Chrischilles et al. (1989)]. Forstrom et al. (1990) and Carter et al. (1984) report similar results; clinical pharmacy services appear to effectively improve prescribing practices.

These studies have indicated that at least some forms of clinical pharmacy services are effective in improving prescribing behavior. At this point, however, the literature in this area has not adequately addressed the effects of clinical pharmacy services on patient outcomes and costs. Furthermore, the nature and scope of the pharmacist's role varied across the studies, making a comparison of results difficult. It is also likely that the effectiveness of clinical pharmacy services varies across different clinical settings. It is not clear, therefore, that all forms of clinical pharmacy services will be effective in improving prescribing practices.

### 1.4.7 The Role of Incentives

It is widely believed that in addition to the educational interventions discussed above, financial and other incentives can be used to modify physician behavior. Although some observational studies suggest that physicians do modify their behavior in response to financial incentives, there are very few randomized trials in this area [Greco and Eisenberg (1993)].

Hillman et al. (1989) find that financial incentives in the context of a managed care organization affect hospitalization rates and office visits. [Hillman et al. (1989), 90] In a pre/post study, Morrow et al. (1995) report that financial incentives were associated with improved preventive health procedures, such as higher rates of childhood immunizations. As the authors state, however, this study's design cannot adequately assess causation issues [Morrow et al. (1995)].

Very few rigorous studies have assessed the impact of financial incentives on physicians' prescribing behavior. Hueston et. al. (1997) compare potentially unnecessary antibiotic prescribing between random samples of patients in a staff model health maintenance organization (HMO) and in a fee-for-service (FFS) environment. These researchers find that even though the HMO physicians faced financial incentives and risks (such as capitation rates, utilization review, and salaries partially linked to HMO profits), HMO physicians were *not* less likely than FFS physicians to prescribe antibiotics for upper respiratory infections and acute bronchitis. Nevertheless, as expected, HMO physicians did have lower rates of diagnostic testing for these illnesses compared to FFS physicians [Hueston et al. (1997)]. Hueston et al. (1997) did not control for those characteristics of patients and physicians who select into HMOs or remain in fee-for-service that may affect the outcomes of interest. Therefore it is not possible to conclude that physician prescribing practices are insensitive to financial incentives.

Because of the dearth of research in this area, it is not clear that financial incentives can improve prescribing behavior. As the use of incentives becomes more widespread, it is important that the effectiveness of this method is rigorously evaluated.

### 1.4.8 The Role of Regulatory and Administrative Restrictions

Regulatory and administrative restrictions are commonly used by hospitals, managed care organizations, and public insurance programs to modify prescribing practices. These restrictions include prescription limits, formularies, category exclusions, and prior authorization systems. These restrictions are designed to contain costs, deter fraud and prevent inappropriate prescribing [Soumerai et al. (1993)].

In particular, as of 1992, Medicaid programs in 34 states had instituted prior authorization systems [Smalley et al. (1995)]. Smalley et al. (1995) report that in Tennessee, a prior authorization program designed to promote Medicaid patients' use of generic nonsteroidal antiinflammatory drugs (NSAIDS) was very effective. Using interrupted time series analysis, these researchers find that two years after the prior authorization policy was implemented, mean annualized expenditures for NSAIDS had declined by 53 percent compared to the baseline year. Smalley et al. (1995) report no significant differences in expenditures for other related drugs or for outpatient services before and after the policy was in effect [Smalley et al. (1995)].

Similarly, Durbin et al. (1981) and Ahluwalia et al. (1996) report that prior authorization policies are effective in the hospital setting. Durbin et al. (1981) evaluate a program that automatically discontinued antibiotic prescriptions at clinically appropriate times unless the prescription was reviewed and reordered by the prescriber. Although no control group was used in this study, the authors report reductions in antibiotic use in the hospital's surgery unit [Durbin et al. (1981)]. In a more recent study, Ahluwalia et al. (1996) find that the use of NSAIDS requisition forms reduced brand name NSAIDS prescriptions by 34 percent. This study had an observational, pre/post study design with no control group [Ahluwalia et al. (1996)].

Payment restrictions also have been used to change prescribing practices. Soumerai et al. (1987) use a time series with comparison design to study the effects of a Medicaid three prescription per month cap in New Hampshire. The prescription limit policy reduced prescriptions by 46 percent among multiple drug users [Soumerai et al. (1987)]. In a later study of the New Hampshire prescription cap, however, Soumerai et al. (1991) note that multiple drug users subjected to this policy were twice as likely as control group patients to enter nursing homes during the study period [Soumerai et al. (1991)].

This important finding suggests that although there is evidence that regulations are effective in changing prescribing behavior, it is not clear how these policies affect patient outcomes. An effective prescription regulation, therefore, may not be cost-effective or ethical. As Soumerai et al. (1993) and others have suggested, it is important that researchers address the first-order effects (i.e. use of drug substitutes) and second order effects (i.e. hospitalizations) before making conclusions about administrative and regulatory drug programs [Soumerai et al. (1993)].

### 1.4.9 Conclusions

The literature reviewed in the previous sections offers evidence that prescriber behavior can be modified through particular forms of interventions. Face to face, individual academic detailing, intensive group education programs, feedback letters, prior authorization and prescription limit policies all demonstrate the potential to effectively change prescribing practices. It appears that the effectiveness of an educational intervention varies by the following characteristics: (1) intensity of effort (i.e. number of educational sessions); (2) professional background of implementors (i.e. physician led intervention vs. pharmacist led intervention); and (3) objective of intervention (i.e. lower costs, improve patient outcomes). Intensive interventions led by physicians and doctors of pharmacy are effective in improving prescribing practices. Cost saving, less intensive interventions led by pharmacists are not as effective.

Nevertheless, even if some of the methods described previously are effective, in most cases there is no evidence that they are cost-effective. As many researchers have suggested, the most cost-effective method of changing prescriber behavior may vary by the prescribing practices of the physician. For a physician with a long history of prescription problems, the benefits of individual academic detailing may outweigh the costs. For a broad group of prescribers, however, a less costly feedback intervention may be more appropriate.

In addition to the lack of information on cost-effectiveness, there is also very little research on patients' long and short run clinical outcomes. At least one group of researchers has found that a regulation intended to change prescriber practices resulted in increased nursing home admissions. Because changes in prescription drug regimens can have important effects on health, it is important that clinical outcomes be assessed in the evaluations of these programs and policies.

The general objective of the interventions described in this review is to find a low-cost, effective method of improving prescribing practices so that patient outcomes are improved and/or drug costs are reduced. At this time, there is a growing body of literature on the effectiveness of these interventions, but cost and outcomes issues have not been addressed adequately. Additionally, the effectiveness of the rapidly spreading computerized feedback system has not been studied. With future research in these areas, cost-effective interventions can be implemented that both improve patient outcomes and reduce health system costs.

# 2.0 Evaluation Research Tools

This chapter describes two research tools used in the evaluation: the PSU screener and a bibliographic database that relates prescribing problems that are codified in PSU screening criteria to adverse clinical outcomes. Both of these tools played an important role in the evaluation. The screener provided a mechanism whereby prescription drug claims records from several states and over several years—encompassing millions of records—could be examined for occurrences of DUR screen failures. The bibliographic database provided a means to assist in the selection of clinical outcomes that would be appropriate for assessing the impact of the demonstration.

## 2.1 The PSU Screener

### 2.1.1 History of the Penn State Screener

*Development of Screening Criteria*

Effective DUR requires a broad collection of different screening criteria that have relevance and validity regarding the appropriateness of drug therapies. This is especially important when DUR is applied to prescribing and dispensing of drugs to the elderly. When the Medicare Catastrophic Coverage Act (MCCA) was passed in 1988 it included a prescription drug benefit and called for a program of drug utilization review to be implemented. The lack of a comprehensive and valid set of DUR criteria appropriate for outpatient DUR prompted researchers at the Center on Drugs and Public Policy at the University of Maryland (UM); (Dr. D. Knapp, Principal Investigator) and the Geriatric Pharmacy Institute at the Philadelphia College of Pharmacy and Science (PCPS); (Dr. G. Erwin, Co-Principal Investigator) to jointly propose a project for a systematic model for developing screening criteria for outpatient DUR for the elderly. This project was funded by HCFA under cooperative agreement No. 17-C-99406: *A Model for Developing Methodological Strategies for Outpatient Drug Use Review Under the Medicare Catastrophic Coverage Act of 1988.*

The original HCFA cooperative agreement with the UM/PCPS team called for development of criteria that were appropriate for an *elderly* outpatient population. The UM/PCPS team developed these elderly-specific criteria during 1990 and presented them to HCFA in their first year progress report in March, 1991. After repeal of the Medicare Catastrophic Coverage Act, HCFA requested that the criteria be modified to be appropriate for non-elderly Medicaid beneficiaries. The UM/PCPS final report to HCFA, issued in December, 1992, reflects the newer focus of the screening criteria as well as other changes that resulted from continued examination and evaluation of the draft screening criteria by the expert panels.

*Development of the PSU Screener*

During the period that the screen criteria were being developed, researchers at the Pennsylvania State University (PSU) were awarded a cooperative agreement to conduct a series of studies relating to prescription drug use by the elderly (Cooperative Agreement 17-C-99392, *The Impact of Prescription Drug Coverage on Aged Medicare Beneficiaries*; Dr. Bruce Stuart and Dr. Frank Ahern, Co-Principal Investigators). One of the studies called for the assessment of the impact of DUR on clinical outcomes associated with misprescribing. The PSU team sought and received permission to develop a computer-assisted screening program based on the UM/PCPS criteria. The earliest version of the PSU Screener was developed using the elderly-specific UM/PCPS criteria and was successfully applied to the prescription claims database of Pennsylvania's Pharmaceutical Assistance Contract for the Elderly (PACE).

### 2.1.2 The PSU Screener: An Overview

A major task for the HCFA DUR Demonstration Evaluation Project involved the construction of an efficient computer-screening program that could accurately identify instances of sub-optimal drug therapy as specified in the UM/PCPS criteria. Although the PSU Screener was originally developed specifically to implement the 1992 version of the UM/PCPS screening criteria, it has been updated to accommodate revisions to the criteria and the addition of new classes of criteria.

The PSU Screener is a knowledge-based (expert) system. It is not intended to replace the tasks of clinical monitors but rather to identify exceptional conditions (screen failures) so that they can be brought to the attention of these specialists. The PSU Screener functions as a research tool for the purpose of conducting epidemiologic and economic studies on potentially inappropriate drug therapy.

The PSU Screener uses the Master Drug Data Base (MDDB) product from Medi-Span of Indianapolis, Indiana as the source of drug product information. MDDB is a comprehensive database of drug information updated on a periodic basis. Penn State has subscribed to the MDDB service since April, 1989 and maintaines the currency of the file with monthly updates. At the beginning of each year the current MDDB master is saved for use in historical analysis. This is necessary because the MDDB file is adjusted over time. In addition to new records, the monthly update contains corrections, deletions, and reassignments of the records contained on the master. National Drug Codes (NDC) corresponding to drug products may be inactivated or even reassigned as time goes on. The PSU Screener is capable of conducting retrospective DUR screening using time-specific NDC numbers. The screener identifies 61 specific criteria organized into eight pharmacotherapeutic categories: Angiotensin-Converting Enzyme (ACE) Inhibitors, Antidepressants, Antipsychotics, Benzodiazepines, Calcium Channel Blockers, Digoxin, Histamine-2 Receptor Antagonists (H2RAs), and Nonsteroidal Anti-inflammatory Drugs (NSAIDs).

The PSU Screener performs retrospective DUR using drug claims data and patient demographic information. At present the PSU Screener performs five types of DUR screens:

1. Over/Under Dosage – Excessive daily dosage computed from the strength and days supply of the drug, or dosage too small for effective therapy computed from strength and days supply.

2. Duration of Therapy – Utilization of a drug or therapeutic class for longer than is recommended. Utilization of the acute dosage level of a drug for longer than is recommended. Utilization of a drug for too short a period for appropriate therapeutic efficacy.

3. Drug-Drug Interactions – A drug having a pharmacokinetic and/or pharmacologic effect upon another drug with consequent clinical effects. Modification of the therapeutic or toxic effect of one drug upon another by the prior or concurrent administration of another drug. Some drug-drug interactions are due to the additive effects of the drugs involved.

4. Age-based Contraindications – The drug or some other feature of the prescription (dosage, duration) is not indicated based on the age of the patient.

5. Duplicative Therapy (Polypharmacy) – Too many therapeutic classes of drugs. Two or more drugs in the same therapeutic class (usually seen as a result of switching to alternate drugs at the beginning of a new therapy but may also be seen as a consistent pattern-repetitive duplicative therapy).

The Penn State DUR Screener does not currently utilize diagnoses, food, alcohol intake, or medical procedures. Its intended use is with large (i.e. statewide) pharmaceutical assistance and Medicaid programs where such information may not be available. The PSU Screener is designed to batch process prescription claims data. Information extracted from claims include, for example,: beneficiary identification, pharmacy identification, prescriber identification, date of service, National Drug Code (NDC), days supply, and amount of drug dispensed.

The PSU Screener produces two basic types of output files that summarize, for optional periods of time: (1) prescription drug utilization at the user level, including total number of prescriptions and number of prescriptions within screening categories; (2) physician prescribing practices, including number of patients in the data set, total number of prescriptions (claims), and screen-based volume of prescriptions and identification of patients and prescriptions that violate the screening criteria.

### 2.1.3 Study Objective Served by the PSU Screener

The PSU Screener plays a central role in the HCFA DUR Demonstration Evaluation. It is used to compute the occurrence of screen failures for a consistent set of DUR screening criteria across the two demonstration states, the co-experimental state, and the comparison state from 1989 through 1996. The failure rates provide valuable descriptive information regarding the incidence of potentially inappropriate prescribing patterns. To be of any value, the failure counts must be produced with a *consistent* screener program.

The PSU Screener is described as a "consistent screener" to emphasize the view that certain tasks require that the same instrument be used to measure screen failures across sites and across time. If different screener algorithms were used to assess the risk of inappropriate prescribing for the same drug in different sites or different times, any variation found might be real or an artifact of instrumentation. The selection of the UM/PCPS core set for this task is based on the fact that these are the criteria specified in both the OBRA DUR regulations and the DUR demonstration RFP. The UM/PCPS core criteria have the further advantage of having been validated by expert panels convened for that purpose. Other criteria used in proprietary screener programs do not share that advantage.

### 2.1.4 Updating and Implementing the UM/PCPS Criteria

The original PSU Screener was developed using DUR criteria developed in December, 1992. Given the dynamic nature of the pharmaceutical industry it was necessary to conduct periodic updating and revision of both the UM/PCPS criteria and the PSU Screener to accommodate release of new drugs and changes in official product labeling from 1993 through 1995. The PSU Screener was updated twice: first to cover the period 1992 through December, 1994, and again to cover the period through December, 1995. The procedures for each round of updating and implementing the PSU Screener comprise two steps: (1) updating the criteria to accommodate new drugs/labeling changes that have occurred since the last update, and (2) conducting appropriate certification procedures to assure the accuracy of the updated Screener programs.

*Updating the PSU Screener*

The approach for updating the criteria consisted of the following activities. First, Dr. W. Gary Erwin, Pharm. D., identified new drugs/labeling changes within the UM/PCPS criteria sets. This was done through ongoing review of appropriate media and official records. Once a candidate change was identified, the following elements were determined:

- Dosage. Sources included the official product labeling, Facts and Comparisons' Newsletter, and Drugdex.
- Drug-Drug Interactions. Drug-Drug interactions noted as "major" in various drug compendia were listed.
- Duration of Therapy. Sources included the official product labeling, existing duration parameters in the DUR screening criteria, and the peer-reviewed medical literature.
- Duplicative Therapy. The peer-reviewed medical literature, existing duplicative therapy parameters, and clinical logic were used to identify duplicative therapy.

The updated and revised criteria were forwarded to Dennis Gingrich, M.D. and Alan Adelman, M.D. for review, comment, and approval. Drs. Erwin, Gingrich, and Adelman resolved any differences of opinion regarding the new criteria. Approved criteria were forwarded to PSU for incorporation into the PSU Screener. The first round of revisions and updating procedures for the UM/PCPS core criteria were used to modify the existing PSU Screener and served as a basis for the certification procedures described below for the revised PSU Screener.

*Reprogramming and Certifying the Accuracy of the Updated PSU Screener*

The DUR screens use information on ingredient, drug strength, dosage (days supply and doses), onset and duration of regimen, and sequence of specific combinations of drugs. For some screens as much as six months of prior drug use is evaluated. The PSU Screener program that encodes the UM/PCPS narrative descriptions of the criteria required development of extensive sets of computer code and thus careful attention must be given to the issue of the accuracy of the final determination of whether or not a specific drug regimen showed evidence of a screen violation. Determining the accuracy of the revised and updated PSU Screener followed from procedures established for the original version of the PSU Screener. This entailed a multi-stage process. The first stage began with modification to the computer program giving attention to the effects of modification on the internal operations of the program. The expert system was designed with the flexibility to adapt to changes in DUR criteria without significant reprogramming. If the required modification was the addition of a new drug, a change in a dosage criterion, or a change in a factual element, this was accomplished by an update to the DUR Screener knowledge base. If the modification required a change to the logic or rules for a criterion, then the screens rules were changed. The effort required to reprogram revised criteria was proportional to the difference between the existing implementation and the modification. Most modifications required only minor effort.

The last stage in determining the accuracy of the PSU Screener involved comparisons of the Screener's output to that of clinical experts. This was the heart of the certification process. Prior to porting the updated Screener to Jen Associates for production runs, all criteria contained in the UM/PCPS DUR system were certified. The approach to certification involves the following steps. First, a large sample of actual Medicaid drug data from the study states was obtained to create a benchmark database. From these data samples Medicaid patient of candidate cases (screen failure) and noncases (no failure) were produced. These candidate profiles (cases and non-cases) were certified by clinicians experienced as DUR reviewers (Drs. Gingrich, Adelman, and Erwin). The clinical monitors indicated whether they agreed or disagreed with the computer's identification of screen failures and non-failures and they gave detailed explanations of their reasons for disagreement with the computer's identification. It is important to note that the clinical monitors did not judge whether they agreed with the importance or validity of the criteria but rather only on whether or not the computer correctly identified cases and non-cases in accordance with the criteria. The PSU Screener was revised and modified based on the clinical monitors' explanations and comments. Disagreements among clinical monitors, if any, were resolved before the computerized criteria were adjusted.

### 2.1.5 Porting and Verification of PSU Screener

JEN Associates had primary responsibility for processing all Medicaid claims using the PSU Screener. Procedures were developed and implemented to assure that the JEN installation performed in an identical fashion to its mate at PSU. The operational portions of the certified Screener (code modules and data files) were ported to JEN. JEN in turn adapted these files to their data management system and recoded them to fit the requirements of large scale processing. Once this was accomplished, the verification process entailed comparison of JEN's outputs from a test data file containing known cases and noncases.

### 2.1.6 Screen Definitions

Each of the DUR screens involves specification of one or more prescription drugs. The exact specifications for the screens, including drug names, dosage levels, interacting drugs, age limits, etc. are provided in the PSU Screener Documentation Manual and the two Update Manuals. A brief description of each screen is provided below in Exhibit 2.1.1.

**Exhibit 2.1.1**

**HCFA DUR Demonstration Evaluation**
**Descriptions of UM/PCPS Criteria used in the PSU Screener**

| Therapeutic Drug Class/ Screen Name | Screen Description and Rational |
|---|---|
| **Angiotensin-Converting Enzyme (ACE) Inhibitors** | |
| ACE Inhibitor w/ Lithium | Concurrent use of ACE inhibitors with lithium may increase serum lithium concentrations and precipitate symptoms of toxicity. |
| ACE Inhibitors, Adult Dosage | The adult maximum daily dose of an ACE inhibitor should not exceed the dose for the relevant drugs listed in the Update Manual. |
| Multiple ACE Inhibitors | Concurrent use of two or more ACE inhibitors is not justified since there is no clinical evidence to suggest that combinations improve outcome. |
| **Anti-Depressant Drugs** | |
| Antidepressant w/ >1 Antimuscarinics or Antispasmodics | The combined use of cyclic antidepressants and two or more antimuscarinic/antispasmodics may result in additive anticholinergic side effects (confusion, hallucination, nightmares). |
| Antidepressant, Adult Dosage | The criteria dose ranges (for amoxapine, clomipramine, and doxepin) reflect the caution of the expert panel, based on clinical considerations. The adult maximum daily dose of a cyclic antidepressant drug should not exceed the dose listed for the relevant drugs listed in the Update Manual. |
| Antidepressant w/ MAO Inhibitor | The concurrent use of antidepressants and monoamine oxidase inhibitors has been shown to induce seizures and hyperpyretic crisis and is not recommended. At least 14 days should elapse between the discontinuation of the MAO and the initiation of an antidepressant. |
| Amoxipine and Antiparkinsonians | The use of antiparkinsonian anticholinergic drugs greater than three months at an accelerating dose after the initiation of amoxapine should be identified because such therapy may exacerbate tardive dyskinesia. |
| Antidepressants with Stimulants | The concurrent use of antidepressants with stimulants, including methylphenidate, amphetamines, and phenmetrazine, is identified. |
| Multiple Antidepressants | The concurrent use of two or more antidepressant drugs, including cyclic antidepressants and monoamine oxidase inhibitors, is identified. |
| Impramine w/ Type I Antiarrythmics | The combined use of imipramine with Type I antiarrythmics should be avoided due to the additive effects on cardiac conduction. |
| Tricyclic Antidepressants w/ Maprotiline with Clonidine | The concurrent administration of tricyclic antidepressants or maprotiline with clonidine should be avoided due to the risk of paradoxical hypertension. |
| Tricyclic Antidepressants or Maprotiline w/ Guanethidine and/or Guanadrel. | The concurrent administration of tricyclic antidepressants or maprotiline with guanethidine and/or guanadrel may result in inhibition of the hypotensive effects of this drug. |

**Exhibit 2.1.1 (continued)**

## HCFA DUR Demonstration Evaluation
## Descriptions of UM/PCPS Criteria Used in the PSU Screener

| | |
|---|---|
| **Anti-Psychotic Drugs** | |
| Antipsychotic Drugs, Adult and Pediatric Dosage | The adult and pediatric maximum daily dose of an antipsychotic drug should not exceed the dose listed in the relevant table in the Update Manual. |
| Anti-Psychotic Agents, Pediatric Contraindication | Some anti-psychotic agents are not recommended within certain pediatric age groups. These are identified in the Update manual. |
| Antipsychotic Drugs, Duration | Continuous antipsychotic therapy at the maximum daily dose for greater than six months without an indication that a dosage decrease was attempted is identified. |
| Anti-Psychotic Agent w/ >2 Antimuscarinics or Antispasmodics | The combined use of any of the antipsychotics and three or more antimuscarinic/antispasmodics may result in additive anticholinergic side effects (confusion, hallucination, nightmares) is identified. |
| Antipsychotic w/ Anti-Parkinsonian | The use of antiparkinsonian anticholinergic drugs for a duration greater than three months after initiation of an antipsychotic should be identified because such therapy may exacerbate tardive dyskinesia. |
| Antipsychotic w/ Levodopa | The use of antipsychotics and levodopa together is not recommended because the phenothiazines may inhibit the antiparkinsonian effect of the levodopa. In addition, levodopa has not been demonstrated to be effective in the treatment of drug-induced extrapyramidal effects. |
| Antipsychotic Drugs, Duplicative Therapy | The concurrent use of two or more antipsychotic drugs is identified. |
| **Benzodiazepines** | |
| Benzodiazepines, Dosage | The adult maximum daily dose of a benzodiazepine should not exceed the maximum listed in for the relevant drugs. |
| Benzodiazepine w/Antihistamine | |
| Benzodiazepines w/ Anxiolytic Drugs | The concurrent use of benzodiazepines with other anxiolytic drugs, including buspirone, meprobamate, diphenhydramine, and hydroxyzine, is not indicated. |
| Benzodiazepines w/ Sedative/Hypnotic Drugs | The concurrent use of benzodiazepines and sedative/hypnotic drugs, including barbiturates, chloral hydrate, ethchlorvynol, methyprylon, ethinamate, and glutethimide is not indicated. |
| Benzodiazepines, Duration | Use of benzodiazepines longer than four months at less than the maximum daily dose or use longer than 30 days at the maximum daily dose is identified. |
| Estazolam, Flurazepam, Temazepam, and Triazolam, Duration. | Estazolam, flurazepam, temazepam, and triazolam are not recommended for long term use because of their risk/benefit profile. The duration of therapy for benzodiazepines used for the treatment of transient insomnia should be 35 days. |
| Long Half-Life Benzo. w/ >65 Years or Long Term Care Facility | Benzodiazepines with long (greater than 24 hours) elimination half-lives have been demonstrated to be associated with an increased risk of hip fractures in individuals over 65 years of age residing in community and in individuals residing in long-term care facilities. |
| Multiple Benzodiazepines | The combination of two or more benzodiazepines with the exception of clonazepam is not justified. Therefore, the concurrent use of two or more benzodiazepines except clonazepam is identified. |

**Exhibit 2.1.1 (continued)**

## HCFA DUR Demonstration Evaluation
## Descriptions of UM/PCPS Criteria Used in the PSU Screener

| | |
|---|---|
| **Calcium Channel Blockers** | |
| Addition of Digoxin to a Maintenance Therapy of Verapamil | Concurrent use of verapamil with digoxin may inhibit both the renal and non-renal elimination of digoxin. Reduction in the dose of digoxin is necessary to reduce the risk of toxicity. If a patient is receiving maintenance therapy with verapamil and digoxin is added, the initial dose of digoxin should be less than or equal to 0.125mg for tablets or the equivalent for other dosage forms. |
| Addition of Verapamil to a Maintenance Therapy of Digoxin | Concurrent use of verapamil with digoxin may inhibit both the renal and non-renal elimination of digoxin. Reduction in the dose of digoxin is necessary to reduce the risk of toxicity. If a patient is receiving maintenance therapy with digoxin at a dosage equal to or greater than 0.25mg per day and verapamil is added to this regimen, then the subsequent prescription for digoxin should reflect a reduction in the digoxin dose by at least 50%. |
| Bepridil, Initial Therapy | Bepridil is approved only as a second-line agent for angina and should never be used as initial therapy. Therefore, therapy with bepridil is marked for review due to its Class I antiarrhythmic properties and its propensity to induce serious arrhythmias. |
| Calcium Channel Blockers, Adult Dosage | The adult maximum daily dose of a calcium channel blocker should not exceed the dose listed in the Update Manual. |
| Calcium Channel Blockers w/ Cyclosporine | Concurrent use of calcium channel blockers with cyclosporine may lead to a rise in the serum concentrations of cyclosporine and serum creatinine. Concurrent use of all calcium channel blockers (except nifedipine) prescribed with cyclosporine is identified. |
| Multiple Calcium Channel Blockers | The combination of two or more calcium channel blockers is not justified. All cases of duplicative calcium channel blocker therapy should be reviewed. The exception to this would be a # 21 day prescription episode of nimodipine (assumed to be for the treatment of subarachnoid hemorrhage) plus any other calcium channel blocker. |
| Nimodipine, Duration | Nimodipine is currently approved only for the treatment of subarachnoid hemorrhage with a duration of therapy not to exceed 21 days. |
| Verapamil w/ Beta-Adrenergic Blocking Agents. | Congestive heart failure, AV conduction disturbances, and sinus bradycardia may be more frequent with concurrent use of a beta-adrenergic blocking agents and verapamil. Verapamil should be avoided in patients with severe left-ventricular systolic dysfunction because of its profound negative inotropic effects, particularly in those patients receiving a beta blocking agent. The combination of verapamil plus any beta blocker is identified. |

**Exhibit 2.1.1 (continued)**

## HCFA DUR Demonstration Evaluation
## Descriptions of UM/PCPS Criteria Used in the PSU Screener

| | |
|---|---|
| Verapamil w/ Disopyramide or Flecainide | Concurrent use of verapamil with disopyramide or flecainide is not recommended. Deaths have been reported because all medications possess negative inotropic properties which are additive when administered together. |
| **Digoxin** | |
| Add Cyclosporine to Digoxin | Severe digoxin toxicity occurred in some patients who were started on cyclosporine. |
| Digoxin + Amiodarone or Quinidine w/o Dosage Drop of Digoxin | Quinidine reduce the non-renal clearance of digoxin as well as the volume of distribution by displacing digoxin from tissue-binding sites. Reduction in the dose of digoxin is necessary to reduce the risk of toxicity. Therefore, in patients maintained on $ 0.25mg of digoxin daily, a reduction in digoxin dosage of at least 50% must occur if amiodarone or quinidine is added to their drug regimen. |
| Digoxin, Adult and Pediatric Dosage | The maximum daily maintenance dose of digoxin per age group should not exceed the dose listed in the Update Manual. |
| Digoxin Added at Higher Than Recommended Dosage to Amiodarone, Quinidine, or Propafenone | If a patient is receiving maintenance therapy with with amiodarone, quinidine, or propafenone and digoxin is added, then the initial daily dosage of digoxin should be # 0.125mg for tablets, or the equivalent for other dosage forms. |
| Multiple Digoxins | The concurrent use of two or more digitalis glycosides is reviewed since there is no evidence to suggest that the combination is superior to single drug therapy . |
| **Histamine $H_2$-Receptor Antagonists** | |
| Cimetidine w/ Phenytoin, Methylxanthines, Warfarin, or Procainamide | In evaluating the cimetidine interactions for drug use review, the order in which therapy was initiated is important. It is reasonable and clinically safe to add a prescription for phenytoin, methylxanthine, warfarin, or procainamide to a patient receiving cimetidine since the dosage of the second drug could be titrated to an objective parameter such as plasma concentration or prothrombin time. However, if a patient is maintained on phenytoin, methyxanthines, warfarin, or procainamide, it would be considered inappropriate to add cimetidine to this drug episode since other effective, non-interacting histamine $H_2$-receptor antagonist therapy is available. |
| $H_2$-Receptor Antagonists & Sucralfate on Same Day | Histamine $H_2$-receptor antagonists and sucralfate should not be initiated on the same date. This would imply that single drug therapy has not been evaluated |

**Exhibit 2.1.1 (continued)**

## HCFA DUR Demonstration Evaluation
## Descriptions of UM/PCPS Criteria Used in the PSU Screener

| | |
|---|---|
| $H_2$-Receptor Agonists w/ Ketoconazole | Since the oral absorption of ketoconazole is dependent on an acidic environment, the combination of any histamine $H_2$-receptor antagonist with oral ketoconazole is considered to be inappropriate. Use of a histamine $H_2$-receptor antagonist results in an increase in gastric pH and limits the absorption of ketoconazole. Therefore, concurrent use of a $H_2$-receptor antagonist with oral ketoconazole is identified |
| $H_2$-Receptor Antagonists w/ Misoprostol, Omeprazole, or Sucralfate | Even though the mechanism of action is different, the use of misoprostol, omeprazole, or sucralfate with any histamine $H_2$-receptor antagonist is considered concurrent and therefore inappropriate. |
| $H_2$-Receptor Antagonists, Acute Duration | Literature and package labelling indicate a maximum of 56 days (eight weeks) for the treatment of acute peptic ulcer disease, gastritis, and duodenitis. Sixty-two days rather than 56 days is used because most prescriptions are dispensed for 30- or 31-day supplies. |
| $H_2$-Receptor Antagonists, Dosage | If the days supply is less than or equal to 62 days, the maximum daily dose criteria for acute therapy should not exceed the dose listed in the Update Manual. If the duration of therapy exceeds 62 days, then dosage criteria for maintenance therapy or GERD is used. Eight weeks is the recommended duration of acute therapy for peptic ulcer disease. |
| $H_2$-Receptor Antagonists, GERD Duration | Full dose therapy for GERD should not exceed 93 days or a three-month supply as defined by the database. Literature and package labelling indicate that the effectiveness of acute dosages beyond 12 weeks for the treatment of gastroesophageal reflux disease has not been thoroughly evaluated. |
| Multiple $H_2$-Receptor Antagonists | The concurrent use of two or more histamine $H_2$-receptor antagonists is reviewed since there is no clinical evidence to suggest that combinations improve outcome. |

**Non-Steroidal Anti-Inflammatory Drugs (NSAIDS)**

| | |
|---|---|
| Ketorolac with Probenecid | Due to a decrease in ketorolac clearance from concurrent probenecid administration, it is contraindicated because of the risk of serious adverse reactions and risk of bleeding. |
| Mefenamic Acid oe Phenylbutazone Duration | Use of mefenamic acid or phenylbutazone is recommended for no longer than seven days because of their ability to cause diarrhea. |
| Multiple Nonsteroidal Anti-Inflammatory Drugs | The combination of two or more NSAIDS is not justified except the use of # 1300mg daily of aspirin plus another NSAID (unfortunately aspirin use is not usually included in a Medicaid database). |

**Exhibit 2.1.1 (continued)**

**HCFA DUR Demonstration Evaluation**
**Descriptions of UM/PCPS Criteria Used in the PSU Screener**

| | |
|---|---|
| NSAID w/ High Dosage of Methotrexate | Concurrent use of an NSAID with high dose methotrexate (greater than 30mg /day) may produce sometimes fatal methotrexate toxicity. This toxicity has been reported to occur when methotrexate is used to treat rheumatoid arthritis. Concurrent use of NSAIDs with high dose methotrexate is identified. |
| NSAIDs w/ Lithium | NSAID induced renal prostaglandin inhibition may reduce renal lithium excretion, leading to elevated serum lithium concentrations and neurotoxicity. Concurrent use of lithium with an NSAID except salicylates is identified. |
| Selected NSAIDs w/ Oral Anticoagulants | The NSAIDs may increase the risk of bleeding in patients receiving oral anticoagulants through several mechanisms, including: (1) displacement of oral anticoagulants from protein-binding sites, (2) inhibition of oral anticoagulant metabolism, (3) inhibition of platelet aggregation, and (4) mucosal injury. Available literature suggests caution with diflunisal, ibuprofen, flurbiprofen, indomethacin, meclofenamate, mefenamic acid, fenoprofen, salicylates, phenybutazone, piroxicam, and sulindac. Concurrent use of an anticoagulant with these NSAIDs is identified. |
| NSAIDS, Adult Dosage | The adult maximum daily dose of an NSAID should not exceed the dose listed in the Update Manual. |
| Phenybutazone, Initial Therapy | Phenylbutazone is not considered by the expert panel to be the drug of first choice. Therefore, unless a prescription for another NSAID is documented prior to initiation of phenylbutazone treatment, use of phenylbutazone is identified |
| Phenylbutazone w/ Oral Antidiabetics | The hypoglycemic effects of oral antidiabetic drugs may be enhanced by phenylbutazone. Concurrent use of oral antidiabetics and phenylbutazone is identified. |
| Salicylates and Probenecid | Large doses of salicylates may inhibit the uricosuric properties of probenecid and such use should be avoided. Concurrent use of salicylates and probenecid is identified. |
| Add Triamterene to Indomethacin | Although rare, the addition of triamterene to a maintenance dose of indomethacin may cause reversible acute renal failure. |

Source: Update Manual, Penn State University DUR Screener

## 2.2 The DUR Outcomes Database

### 2.2.1 Development of *The DUR Outcomes Bibliographic Database*

A significant component of the overall demonstration is the evaluation of the impact of the state-specific DUR programs on clinical outcomes that may be associated with prescribing problems. Selection of the relevant clinical outcomes is first step toward this goal. There is no central database or reference source containing information that comprehensively and definitively links DUR screening failures to expected clinical outcomes. However, there does exist an extensive scientific and clinical literature that documents adverse clinical effects due to use, misuse, and misprescribing of prescription drugs. This literature is diverse and scattered, including case reports in clinical journals, summary reports in drug compendia, and highly controlled scientific studies in journals. In order to examine the impact of the DUR interventions on clinical outcomes it was necessary to develop a mechanism to systematically review and evaluate the literature on clinical outcomes associated with failure of DUR criteria. The primary purpose of the bibliographic review was to obtain a candidate set of clinical outcomes that could then be further reviewed by a sub-committee of clinicians and researchers. The selection and review of candidate outcomes was facilitated by the development of an outcomes/screener database—more formally called *The DUR Outcomes Bibliographic Database*. This database was then used to assist in the final selection of clinical outcomes for the five outcomes studies discussed in later chapters of this final report.

In a prior HCFA cooperative agreement, the PSU research team developed a set of prototypical procedures that identified clinical outcomes that were applicable to the earliest draft version of the UM/PCPS criteria. These procedures were further developed and were proposed as a model for selecting clinical outcomes for the current DUR Demonstration and were implemented after extensive review and discussion by the Technical Advisory Group. The following presents a summary of the attributes of the database.

*Attributes of The DUR Outcomes Bibliographic Database*

In the initial planning of the database it was decided that a computerized relational bibliographic database would be essential to the evaluation of the literature review. Such a database could be initially built upon research on prescription drugs and DUR rules contained in the UM/PCPS screening criteria, and then be updated as needed to accommodate future criteria and newly published results relating to clinical outcomes and DUR criteria. Several computer software products were evaluated for use as the basis for the database. Parodox--an easily available and inexpensive relational database was selected for use.

*Selection of Candidate Clinical Outcomes*

The first task was to identify potential or candidate clinical outcomes that were relevant to the UM/PCPS DUR screens (i.e., the screens implemented in the PSU screener. A variety of key terms were developed to initiate the search strategy, including for example:

- Pharmacotherapeutic classes (e.g., *benzodiazepines, antidepressants*) underlying the organization of the UM/PCPS screens;

- Pharmacotherapeutic classes named within the screening criteria (e.g., benzodiazepines with *antihistamines* or *other anxiolytics*;

- Specific drugs named in the UM/PCPS criteria (e.g., ACE inhibitor with *levadopa*) or contained in the drug classes (e.g., *estazolam, triazolam*);

- Specific outcomes named in the UM/PCPS screening criteria as part of the rationale for specific screens (e.g. *falls* associated with use of long half-life benzodiazepines; *congestive heart failure* associated with use of verapimil and beta- adrenergic blocking agents);

- Author lists from published and unpublished research papers that document clinical outcomes.

Using the dictionary of key words, the most recent five years of primary literature was searched using electronic bibliographic databases (e.g., Medline, Psych-Info). Indexes of selected secondary sources were manually searched including the major drug compendia (American Hospital Formulary Service, American Medical Association Drug Evaluations, United States Pharmocopoea), and texts related to the therapeutic and problematic use of drugs (e.g., Gilman et al., 1990; Gitlin, 1990). The available key words in the database are grouped into several categories. For example, a search may be conducted using either "antianginals" or "antihypertensives" as key words. Another search strategy utilizing this database can be conducted using the pharmacologic group of the drug(s) such as "tricyclic antidepressants" or "histamine$_2$-receptor antagonists". Outcomes related to any of the UM/PCPS screens can also be used as key words. Searches can be conducted using "congestive heart failure" or "dizziness and giddiness" , among others, as key words. The terminology used to describe the outcomes matches, as closely as possible, that of the ICD-9-CM (4th revision, 1994) manual. Therefore a search based on a specific outcome can be conducted using either the ICD-9-CM code(s) or a descriptor of that outcome. The outcomes are subset into two categories — one contains those outcomes which are listed in the ICD-9-CM manual as a single code. The other category contains those outcomes that are described by a series of ICD-9-CM codes.

Because a reference may have several relevant key words associated with it, the bibliographic citation provided by the database lists all appropriate key words. For example, a search based on the key word "hypotension" may yield a list of several references. If one of those references also lists "lisinopril" as a key word, this means that an independent search using "lisinopril" as a key word would also have yielded this particular reference. Similarly, the references in the database are cross-referenced by screen. For example, if a search is conducted to obtain all reference pertaining to screen involving verapamil, the citations provided will also list any other screens that also involve the drug verapamil. It is also possible that the database can be used to conduct searches based on more than one key word grouped together.

The process of linking medical outcomes to the UM/PCPS DUR criterion involved a meta analysis of existing primary and secondary pharmaceutical and clinical literature. The initial phase of the meta analysis consisted of a search through secondary sources in order to link specific medical outcomes to the screen criterion. The next phase consisted of a literature review of primary sources using the MedLine database. The final phase consisted of linking the specified outcomes to their appropriate

ICD-9-CM codes. For each candidate outcome condition identified the most specific ICD-9-CM diagnostic code were identified. For some conditions it is possible to define a series of related ICD-9-CM codes that encompass variations of the diagnostic condition in question. Also included are all external cause-of-injury codes (E-codes) that specified poisonings (not purposely inflicted) due to drugs and medicinal agents (e.g., E854.0, accidental poisoning by antidepressants) and E-codes for adverse effects of drugs in therapeutic use (e.g., E939.0 adverse effects from antidepressants).

### 2.2.2 Selection of Candidate Clinical Outcomes

*Development of Summary Reports*

The Penn State team coordinated the task of selecting candidate clinical outcomes. The *DUR Outcomes Bibliographic Database* was used extensively in the selection and evaluation process. This was done by developing various summary reports on the evidence linking clinical outcomes with specific DUR criteria. The summary reports took the form of short meta-analyses [(Dickersin and Berlin (1992)] and included quantitative and qualitative summaries of the literature for each candidate outcome. A typical summary report contained the following elements:

1. Reference to the source material for the descriptions of the screening criteria: (a) the UM/PCPS Final Report; and (b) the PSU Screener Documentation Report. These references provided the exact definitions of the screens and the details relating to the screening parameters (exact dosages, durations, interactive drugs, special conditions, etc.).

2. Summary of the relevant literature. The report summarized the literature that supported the association between a specific outcome or class of outcomes with one or more of the DUR screens. When possible, the summary included information on the relative strength of the association, the presence of dose-response relationships, and information on the study populations to which the study refers.

3. Identification of the relevant ICD-9CM codes and coding groups including excluded codes, if any.

4. A summary table of the proposed clinical outcomes and screening criteria. This provided an overview of the clinical outcomes, relevant ICD-9CM codes, DUR screens, and other information that might have be relevant to selection of the clinical outcome.

5. References. The bibliographic information for the studies cited in the report.

*Review of Candidate Clinical Outcomes*

Reports developed from the of the *DUR Outcomes Bibliographic Database* were mailed to the members of the Clinical Outcomes Subcommittee of the full Technical Advisory Group. The Subcommittee members separately reviewed the reports and provided comments and their clinical opinions back to the Penn State group. On June 5th, 1995 a report was made to the full TAG on the status of the selection process. Preliminary results of the Subcommittee's reviews of the candidate outcomes were presented for discussion. Additional data were also presented on outcome prevalence among Iowa and Washington Medicaid recipients. The discussion of the TAG was focused mainly on the need to further reduce the set of eligible candidate outcomes and on the need to identify a reasonable number of outcomes that could be unequivocally evaluated as a consequence of the intervention.

On October 12, 1995 the Clinical Outcomes Subcommittee met in Harrisburg to further discuss the selection of clinical outcomes to be included in the outcomes analyses. The primary goal was to select a reduced set of outcomes for which there is strong evidence that an effective intervention should lead to a change in incidence and for which there is sufficient power to detect a difference if one exists. Selection of eligible candidate clinical outcomes therefore depended not only on the weight of the evidence (supporting research and clinical review) but also on the availability of sufficient sample sizes to warrant inclusion into the outcomes analyses. In order to help make the decision regarding available sample sizes preliminary analyses were conducted on the Medicaid database for samples drawn from the states that are selected for DUR interventions and controls. A one-year pre-demonstration sample of Iowa Medicaid data was obtained for estimation of whether sufficient samples would be available. For the power analyses, the procedures and formula provided by Strom (1989) were used to estimate the required sample sizes to achieve specified levels of power for cohort studies. The results of the preliminary power analyses were provided to the Clinical Outcomes Sub-Committee. The Subcommittee made set of recommendations of clinical outcomes that were forwarded to the TAG and to the analytic teams. These recommendations serve as the basis for the clinical outcomes analyses provided in Chapters 4 and 5.

# 3.0 Drug Use and Prescribing Problems in Four State Medicaid Programs

## 3.1 Background

Recent studies suggest that medication-related problems are a serious and growing public health issue in the United States. In a widely quoted article, Johnson and Bootman (1995) estimate that drug-related morbidity and mortality account for nearly $77 billion per year in added health care expenditures. Bates, Spell, and Cullen (1997) calculate the annual cost of treating adverse drug reactions (ADRs) among hospitalized patients nationally at more than $4 billion. Bootman, Harrison, and Cox (1997) place a similar $4 billion price tag on the cost of treating medication-related problems in nursing homes. A meta-analysis of 39 prospective studies conducted over the past 20 years of the incidence of ADRs in hospitals led Lazarou, Pomeranz, and Corey (1998) to conclude that 2.2 million hospitalized patients suffered serious ADRs in 1994, and 106,000 died because of them—making ADRs the fifth leading cause of death for hospitalized patients in that year. A recent Alliance for Aging Research Report (1998) lists medication-related problems as the fifth most costly "disease" affecting the elderly at $84.6 billion, following diabetes ($92 billion) but ahead of musculoskeletal conditions ($64.8 billion).

These are staggering statistics. Indeed, the numbers are so huge that they engender skepticism about the true extent of the problem. Adding to the skepticism is the fact none of the estimates quoted above is based on a representative empirical database. The Johnson and Bootman (1995) and Bootman, Harrison, and Cox (1997) studies are based on the judgements of expert clinical panels combined with unit cost data from Medicare reimbursement rates. The Bates, Spell, and Cullen (1997) and Lazarou, Pomeranz, and Corey (1998) articles extrapolate ADR rates from small samples of mainly teaching hospitals to the country at large with no adjustments for differences in hospital or patient characteristics. The Alliance for Aging Research estimate is apparently derived from Johnson and Bootman's figures, but attributes the entire cost of ADRs nationally to persons aged 65 and older.

The lack of credible, population-based data on medication-related problems makes it difficult to formulate appropriate policies to combat them. One data source that could potentially provide such information has been largely overlooked. All pharmacies are required by law to perform drug utilization review and the vast majority have computerized DUR programs that review each prescription before it is filled for an array of potential problems. These include drug-drug interactions, age and pregnancy alerts, proper dosage, early refills, and many other concerns. Because these DUR programs maintain central repositories of prescription claims information, they can be programmed to produce exceptions and drug-problem reports at various levels of aggregation. But there are also serious drawbacks to relying on such systems for population-level data on drug-related problems. The systems currently in use are operated by proprietary vendors and even when the DUR criteria they employ are in the public domain, the computer algorithms that implement the screening criteria are not. Furthermore the systems have not be independently validated nor cross-validated with one another. Lastly, the DUR vendors do not publish their findings.

Unlike proprietary DUR screening systems, The PSU screener offers an open architecture with both the DUR criteria and programming conventions available for public inspection. That system—described in depth in the previous chapter—was used to compute population-level prevalence rates for 61 specific drug-related problems occurring in four state Medicaid programs (Georgia, Maryland, Iowa and Washington) from 1989 through 1996. This screening project represents the first application of a consistent DUR screener program to Medicaid data across states. It also represents the first time that longitudinal data on DUR failure rates have been published for these Medicaid populations. In addition to the intrinsic value of this information to policy makers in the study states, the screening procedures themselves can be readily applied in other states.

## 3.2 Methods

The 61 DUR criteria screened for this study were originally developed by expert clinical panels convened by the University of Maryland and the Philadelphia College of Pharmacy and Science (UM/PCPS) under a cooperative agreement with HCFA [(Knapp and Erwin (1992)]. A January 1996 update of these criteria was used to produce the findings reported here. The screening process involved a claim-by-claim review of all Medicaid prescriptions filled in the study states over the eight year period for the following drug categories: ACE inhibitors, antidepressants, antipsychotics, benzodiazepines, calcium channel blockers, digoxin, histamine-2 receptor antagonists (H2RAs) and nonsteroidal anti-inflammatory drugs (NSAIDs). The drug problems screened include overdosing, excess duration of treatment, drug-drug interactions, and duplicative therapy.

The output from the screening process comprised: (1) annual counts of all prescription claims for drug products reviewed by the screener, (2) annual counts of persons using these drugs, (3) annual counts of claims failing each of the 61 screening criteria, and (4) counts of persons with one or more screen failures per year by criterion.

These data were aggregated to produce annual DUR failure rates at both the claim and the person level by year and state. For Maryland, Iowa, and Washington, failure rates were computed for the entire eight-year period, 1989-1996. Georgia presented a problem because the state did not require that pharmacists include days supply on prescription bills until the middle of 1993. Because the days supply is a necessary element to run the screening program, Georgia DUR screening results are only reported for 1994 through 1996. The claim-level rates were computed by dividing the number of screen failures for each DUR criterion by the total number of claims screened for that criterion. Person-level screen failure rates were calculated by taking the ratio of the number of persons failing a specific criterion at least once in the year to the total number of persons using drugs in that criterion set in that same year.

The reason for producing two sets of rates is that each addresses different policy questions. Claims-based failure rates provide important information to DUR systems designers. Since all prospective DUR systems generate alerts at the claim level, this is a logical place to test system performance. DUR screens that generate "too many" alerts may be misrepresenting common therapeutic interventions as problematic when they are not (false positives). Screens that generate few hits may be too tightly configured, missing the true extent of the underlying problem (poor sensitivity). Also, it is generally believed that pharmacists develop patterned responses to DUR screen alerts based on the frequency of the alert, the type of drug problem, and the source of the prescription [(Armstrong and Denemark (1998)]. These issues, too, are best examined at the claim level.

DUR screening at the person level gets to the heart of why online prospective DUR systems were developed—to alert pharmacists so that they can intervene and correct potential drug-related problems before they occur. If DUR programs are successful in this regard, the evidence should be observable at the person level. Because the risk of drug misadventures is not evenly distributed across the entire population, person-based screening should be targeted to known (or suspected) high-risk groups. For this project, DUR screening was performed for the entire Medicaid population in each state and for three high-risk subgroups—community-dwelling elderly,[1] adult diabetics, and multiple drug users. The community-dwelling elderly subgroup was defined as Medicaid recipients aged 65 or older with no nursing home claims during the year. Adult diabetics were defined as people 18 years of age or older with a diagnosis of diabetes (ICD-9 250.xx) and/or a prescription for an oral hyperglycemic agent. Multiple drug users were defined as recipients filling prescriptions in three or more different therapeutic classes during the year.

The DUR screening results are provided in their entirety in Appendix A. This appendix contains 64 four-page tables, 16 for each state, ordered in the follow way: 4 tables present DUR screen failure rates per 100 Medicaid recipients by population subset (all recipients, multiple drug users, adult diabetics, and community-dwelling elderly); 4 tables present counts of persons filling prescriptions for screener-reviewed medications (the denominators for the person-level screen failure rates) by population subset; 4 tables present DUR failures per 100 Medicaid claims by population subset, and finally; 4 tables present prescription counts for screener-reviewed drugs (denominators for the claim-level screen failure rates) by population subset. Each table is organized in the same way. Row headers identify the eight therapeutic drug groups reviewed by the PSU screener. Under each drug group heading is a list of screening criteria applied to the drugs in that group. Column heads identify the calendar years for which the contents of the table apply (eight years for Maryland, Iowa, and Washington; three years for Georgia).

The sheer volume of data contained in these tables makes comprehension difficult, to say nothing about interpreting the information presented. The number of persons reviewed per screening criterion/state/year varies from zero to more than 366,000. The number of claims reviewed ranges from zero to 946,000. The screen failure rates encompass the entire spectrum from 0.00% to 100%. Typical cell sizes for the person-based statistics range between 10,000 to 50,000 recipients screened, with the numbers gradually increasing from 1989 until 1995 after which there are slight declines. Cell sizes in the prescription claims tables typically range between 50,000 and 200,000 claims reviewed per screening criterion per year.

To help make sense of these data, the following sections focus on major comparative patterns evident in drug utilization and screen failure rates: low versus high, rising versus falling, differences among the states, differences among risk groups, variation by drug class and type of drug problem. The analysis is purely descriptive. No statistical tests of significance are performed.[2] No multivariate models are estimated. More attention is devoted to findings at the person level than at the claims level because that is where the potential health consequences are most evident.

---

[1] The focus on community-dwelling rather than institutionalized elderly was driven by the fact that HCFA DUR demonstrations did not include nursing home residents.

## 3.3 Drug Use by Therapeutic Class

Exhibit 3.1 presents state-level summary data on drug use in each of the eight therapeutic categories reviewed for 1989 (1994 for Georgia) and 1996. The top panel of this exhibit shows the number of Medicaid recipients with one or more prescription fills in the drug category for the year. The bottom panel shows the average annual number of fills for persons using products in the drug class. For the most part, recipient counts reflect both the relative sizes of the four state Medicaid programs (Iowa being the smallest and Georgia the largest) and the growth in Medicaid populations over time. More Medicaid recipients use NSAIDs than use drugs in any of the other classes. The drug class with the fewest users is digoxin (except in Georgia where there were fewer benzodiazepine users).[3] The most dramatic growth in recipient counts from 1989 to 1996 occurred in the ACE inhibitor class. The number of ACE inhibitor users more than doubled in Iowa and Washington and tripled in Maryland over this period.[4] In only one instance—digoxin users in Washington—were there fewer recipients in 1996 compared to 1989.

The average prescription counts per user by drug class shown in the lower panel of Exhibit 3.1 indicate very different drug utilization patterns among the four states. Some of the variation is undoubtedly due to differences in the composition of the Medicaid populations, some to state Medicaid policies, and some to physician prescribing behaviors. Overall, utilization rates among users of these eight drug groups are highest in Iowa and lowest in Georgia. The difference between these two states is quite dramatic. The average number of scripts per person/year in Iowa during 1996 was higher in every class—50 percent higher in the case of antipsychotics, benzodiazepines, and NSAIDs. Not unexpectedly, antipsychotics exhibited the highest mean number of fills per year in any drug class. In Iowa and Washington the average recipient of antipsychotic medications filled 11 claims in 1996 or nearly one per month. Although more Medicaid recipients use NSAIDs than any other class of drug, the average level of NSAID use is low—between 2 and 3 prescription fills per person/year across the states.

The table shows a definite rising trend is drug usage. The greatest gains were recorded in the antipsychotic, ACE inhibitor, and calcium channel blocker categories. Washington far outpaced the other states in term of rising numbers of claims per user. Between 1989 and 1996, annual usage rates for antipsychotics rose from 6.9 to 11.1 scripts, a 61 percent increase. Digoxin use increased by 59 percent, ACE inhibitors by 31 percent, and calcium channel blockers by 28 percent. There were a few

---

[2] Strictly speaking, all differences in rates are "statistically significant" since the data represent population values rather than sample values. No attempt is made to generalize the results to other state Medicaid programs or to non-Medicaid populations.

[3] See section on benzodiazepines on page 3-10.

[4] Over this time period ACE inhibitors became first-line therapy for high blood pressure among diabetics and were also more commonly used for other hypertensive patients.

**Exhibit 3.1**

**HCFA DUR Demonstration Evaluation**
**Use of Prescription Drugs in Eight Therapeutic Categories**
**In the Four Study States, 1989-1996**

| | Number of Recipients with Prescriptions by Drug Class, State, and Year | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Georgia | | Iowa | | Maryland | | Washington | |
| **Drug Category** | **1994** | **1996** | **1989** | **1996** | **1989** | **1996** | **1989** | **1996** |
| Ace inhibitors | 41773 | 48257 | 4436 | 11286 | 7533 | 24458 | 8040 | 19272 |
| Antidepressants | 51023 | 51818 | 12947 | 19684 | 10955 | 22244 | 25511 | 33492 |
| Antipsychotics | 40940 | 42965 | 11453 | 14513 | 13160 | 23855 | 18875 | 21932 |
| Benzodiazepines | 8203 | 8303 | 14573 | 18788 | 19673 | 29136 | 20193 | 28920 |
| Calcium Channel Blockers | 64673 | 66895 | 6928 | 13585 | 11782 | 33930 | 10641 | 21404 |
| Digoxin | 28347 | 26961 | 6855 | 7944 | 5986 | 10988 | 10416 | 9984 |
| H2RAs | 71730 | 87172 | 12284 | 20069 | 16882 | 34228 | 28547 | 34276 |
| NSAIDs | 231872 | 207807 | 30953 | 43430 | 39455 | 63645 | 62187 | 58201 |

| | Average Prescription Claims per User in Drug Class, State, and Year | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Georgia | | Iowa | | Maryland | | Washington | |
| **Drug Category** | **1994** | **1996** | **1989** | **1996** | **1989** | **1996** | **1989** | **1996** |
| Ace inhibitors | 6.4 | 6.8 | 6.7 | 7.7 | 5.3 | 5.7 | 5.5 | 7.2 |
| Antidepressants | 4.8 | 4.9 | 6.5 | 6.4 | 5.9 | 5.5 | 5.4 | 6.0 |
| Antipsychotics | 6.6 | 7.1 | 9.0 | 11.0 | 7.8 | 9.3 | 6.9 | 11.1 |
| Benzodiazepines | 4.2 | 4.5 | 6.9 | 6.9 | 6.2 | 6.2 | 5.0 | 5.7 |
| Calcium Channel Blockers | 6.8 | 7.2 | 7.0 | 7.9 | 5.8 | 6.0 | 5.7 | 7.3 |
| Digoxin | 6.9 | 7.1 | 8.1 | 8.4 | 6.4 | 6.2 | 4.1 | 6.5 |
| H2RAs | 2.8 | 4.5 | 5.0 | 5.3 | 4.3 | 4.7 | 4.4 | 5.0 |
| NSAIDs | 2.4 | 2.4 | 3.7 | 3.1 | 3.2 | 2.7 | 3.1 | 3.3 |

*Source:* Medicaid drug data from Georgia, Iowa, Maryland, and Washington, 1989, 1994 and 1996.

instances of declining average utilization rates over the eight-year period: antidepressants and NSAIDs in Iowa and Maryland, and digoxins in Maryland.

## 3.4 DUR Screen Failure Results

### 3.4.1 DUR Criteria Associated with Consistently Low Failure Rates

Low DUR failure rates are to be applauded because they indicate either that practice conforms to expert opinion or that the problems themselves are rare.[5] What constitutes low prevalence in screen failure rates? For three of the 61 criteria screened, there were no instances of DUR failures in any state/year/population subgroup. These three are all drug-drug interactions rated as having class 1 significance in Hansten's Interactions and Updates: antidepressants with an MAO inhibitor, amoxapine with an antiparkinsonian agent, and ketorolac with probenecid. Only in the latter instance were the number or claims or persons screened low in absolute numbers. The number of Medicaid recipients using either ketrolac or probenecid ranged from a low of 373 in Iowa (1989) to a high of 1,727 in Washington (1991). Even rarer is use of nimodipine. Screening for nimodipine duration of therapy identified just a handful of users (20 recipients in Georgia in 1996) and even fewer instances of DUR failure over the entire panel of states and years. By contrast, between 12,000 recipients (Maryland, 1989) and 80,000 recipients (Georgia, 1996) were screened for the antidepressant w/MAO inhibitor interaction. Finding none in such a large population base certainly confirms that practitioners are aware of the problem and abide by the specified drug use criteria.

---

[5] This assumes, of course, that the DUR criteria are valid and have high specificity. The specificity of all PSU screening criteria is high (see discussion in Chapter 2).

Raising the threshold of "low" to less than one screen failure in 400 claims or persons reviewed (< 0.25% failure rate) in any state/year resulted in 13 of the 61 DUR criteria qualifying for low failure rate status (Exhibit 3.2). All but two of these 13 criteria (use of bepridil or phenylbutazone as first-line therapy) reference drug-drug interactions, and all can produce severe consequences. These drug problems cluster in three drug groups: anti-depressants, NSAIDs, and digoxins (by contrast, none of DUR criteria in the ACE inhibitor, benzodiazepine, or H2RA drug groups met this standard for low prevalence of screen failures). As one would expect there is a close match between failure rates at the claim and person levels. In every case but one (the interaction of phenylbutazone with oral antidiabetics), low prevalence of claims failures is associated with low prevalence in failures rates at the person level. And even for that criterion, the highest recorded person-level prevalence was 0.39 percent (Iowa, 1989).

The ketorolac w/probenecid interaction remains the only low-failure-rate criterion in which the number of persons and claims screened was itself small. At least 4,000 recipients and 30,000 claims were screened in each state/year for the other 12 criteria (the typical range was between 10,000 and 40,000 recipients and between 70,000 and 400,000 claims screened). Here again, with such a large risk exposure, the fact that there were so few DUR failures is very encouraging. One reason for this may be that certain criteria, such as cyclosporine added to digoxin or NSAID with high dose methotrexate, are associated with drugs generally prescribed by specialists, which implies controlled use. Other criteria, such as those with digoxin, represent a very narrow therapeutic index that generally encourages careful prescribing by all clinicians.

### 3.4.2 DUR Criteria Associated with Consistently High Failure Rates

At the other end of the spectrum are DUR criteria that are frequently violated. Experts may disagree about what constitutes a "high" DUR failure rate, but few would take issue with the 10 criteria shown in Exhibit 3.3. To be selected for this exclusive list, the failure rate had to exceed 8.0 percent of all claims or persons screened in every state and year of the study. Even with this stringent standard for inclusion, the lower-bound failure rates ranged from 8.1 percent (the concurrent use of an H2RA with misoprostol, omeprazole, or sucralfate) to 53.4 percent (excessive duration of therapy for estazolam, flurazepam, temazepam, or triazolam).[6] At the upper bound, failure rates approached 70 percent (excessive duration of therapy with mefenamic acid or phenylbutazone).

---

[6] The duration of therapy screens for benzodiazepines are complex. The duration of therapy for estazolam, flurazepam, temazepan, or triazolam is set at 35 days whereas the general benzodiazepine duration of therapy screen for all drugs in this class (including the four just names) is dose-related. It is thus possible that a claim may fire both screens simultaneously.

Exhibit 3.2

### HCFA DUR Demonstration Evaluation
### DUR Criteria Associated with Consistently Low Failure Rates in the Four Study States, 1989-1996 (% claims/persons failing criteria per year)

| DUR Criterion | Range of failure rates per claims (%) | Range of failure rates per person (%) |
|---|---|---|
| **ANTIDEPRESSANTS** | | |
| Antidepressants w/MAO inhibitor | 0.00 - 0.00 | 0.00 - 0.00 |
| Amoxapine w/Antiparkinsonian | 0.00 - 0.00 | 0.00 - 0.00 |
| Tricyclic Antidepressant or Maprotiline with Guanethidine/Guanadrel | 0.00 - 0.11 | 0.01 - 0.07 |
| **ANTIPSYCHOTICS** | | |
| Antipsychotic w/>3 Antimuscarinics or Antispasmodics | 0.00 - 0.06 | 0.00 - 0.04 |
| **CALCIUM CHANNEL BLOCKERS** | | |
| Add Verapamil to Digoxin | 0.01 - 0.09 | 0.00 - 0.16 |
| Bepridil: Initial Therapy | 0.00 - 0.00 | 0.00 - 0.03 |
| **DIGOXINS** | | |
| Cyclosporine Added to Digoxin | 0.00 - 0.02 | 0.00 - 0.09 |
| Digoxin + Amiodarone/Quinidine w/o Dosage Drop | 0.00 - 0.05 | 0.01 - 0.10 |
| Multiple Digoxins | 0.00 - 0.09 | 0.00 - 0.04 |
| **NSAIDs** | | |
| Ketorolac w/Probenecid | 0.00 - 0.00 | 0.00 - 0.00 |
| NSAID w/High Dose Methotrexate | 0.00 - 0.01 | 0.00 - 0.01 |
| Phenylbutazone: Initial Therapy | 0.00 - 0.07 | 0.00 - 0.25 |
| Phenylbutazone w/Oral Antidiabetic | 0.00 - 0.13 | |

*Source:* Medicaid drug data from Georgia, Iowa, Maryland, Washington, 1989-1996.
*Note:* "Consistently low failure rates" indicate DUR criteria for which less than 0.25% of claims or persons reviewed failed the criterion in question in every state/year considered.

**Exhibit 3.3**

**HCFA DUR Demonstration Evaluation**
**DUR Criteria Associated with Consistently High Failure Rates in the Four Study States, 1989-1996**
**(% claims/persons failing criteria per year)**

| DUR Criterion | Range of failure rates per claims (%) | Range of failure rates per person (%) |
|---|---|---|
| **ANTIPSYCHOTICS** | | |
| Multiple Antipsychotics | | 9.9 - 16.3 |
| **ANTIDEPRESSANTS** | | |
| Antidepressant Dosage | 11.8 - 21.5 | 14.3 - 26.6 |
| **BENZODIAZEPINES** | | |
| Benzodiazepine w/Anxiolytic | 8.8 - 16.0 | |
| Benzodiazepine: Duration | 27.4 - 48.6 | 26.0 - 41.4 |
| Estazolam, Flurazepam, Temazepam, Triazolam: Duration | 53.4 - 66.4 | 15.6 - 56.0 |
| Long Half-life Benzodiazepine: >Age 65 | 21.8 - 39.9 | 20.6 - 37.2 |
| **H2RA's** | | |
| H2RA: Adult Dosage | 11.9 - 35.5 | 9.8 - 39.6 |
| H2RA w/Misoprostol, Omeprazole, Sucralfate | | 8.1 - 14.7 |
| **NSAIDs** | | |
| Mefenamic Acid or Phenylbutazone: Duration | 29.4 - 70.0 | 30.0 - 68.3 |
| NSAIDs, Dosage | | 11.2 - 22.2 |

*Source:* Medicaid drug data from Georgia, Iowa, Maryland, Washington, 1989-1996.
*Note:* "Consistently high failure rates" indicate DUR criteria for which more than 8.0% of claims or persons fail the listed criterion in every state/year considered.

High failure rates do not necessarily mean that large numbers of persons failed the screening criteria. For example, fewer than 150 individuals failed the mefenamic acid or phenylbutazone duration screen in any state or year. Fewer than 300 failed the estazolam, flurazepam, temazepam, or triazolam duration screen in any of the three years of data available from Georgia. On the other hand, between 7,000 and 15,000 Medicaid recipients in Iowa, Washington, and Maryland failed the benzodiazepine duration of therapy criterion each year of the study. Large numbers of individuals also failed the NSAID and H2RA dosage screens. In terms of risk exposure, the number or recipients screened for these high failure rate criteria, ranged from 10 in Iowa for the mefenamic acid/phenylbutazone duration screen (1995) to 250,747 in Georgia for NSAIDs adult dosage (1995).

What might account for such extreme values? Four of the ten criteria associated with high failure rates relate to benzodiazepine use. Seven are associated with high dosage or excessive duration of therapy criteria. The benzodiazepine drug class presents an interesting case since these drugs have been subject to intense scrutiny in the scientific literature, especially concerning misuse of these medications among the elderly. Duration and dosage are types of criteria particularly prone to either subjective clinical judgement or high false positive rates. Physicians may feel more comfortable in intentionally prescribing outside usual dosage or duration parameters than they would, say, in prescribing drugs with known, serious interactions. It is also possible that some of the reported DUR problems in Exhibit 3.3 are inflated because the computer interpretations of the clinical criteria are insensitive to common prescribing and drug use behaviors (e.g., early refills are flagged as high dosage or a treatment change from monotherapy to polytherapy is flagged as duplicative therapy). Whether the problems are real or artifactual, they clearly should be addressed if DUR screening is to be a useful quality improvement tool.

### 3.4.3 Comparative Findings by Drug Class

The next eight sections examine trends in DUR failure rates for each therapeutic drug class screened with the PSU screener. These sections highlight particularly notable interstate differences and identify drug-related problems that either show significant improvement over time or imply that the problems are becoming worse. The focus here is on person-level rates except where trends in claims-based failure rates diverge significantly from the person-level rates. For ease of interpretation, summaries of the DUR criteria are provided for each therapeutic class.

***ACE Inhibitors.*** There are three DUR screening criteria for the angiotensin-converting enzyme (ACE) inhibitors: maximum daily dose, concurrent use of multiple agents, and a drug interaction with lithium. Failure rates for all three criteria are relatively low across the entire state/year panel (although not low enough to make the 0.25% failure-rate cutoff to warrant inclusion in Exhibit 3.2). The rates range from 0.1 percent to 5.5 percent for the Medicaid populations as a whole and up to 8.3 percent for the high-risk population groups (the rate for multiple ACE inhibitor use among adult diabetics in Georgia, 1996). The prevalence of screen failures for ACE inhibitor dose and duplicative therapy is two to three times higher among community-dwelling elderly than for the entire Medicaid population. Interestingly, failure rates for the ACE/lithium interaction are much less common among the elderly (e.g. a 0.13 percent failure rate for elderly Iowa Medicaid recipients in 1996 compared to 1.78 percent for all Medicaid recipients in the state for that year). There are no consistent trends across the three criteria: the prevalence of ACE/lithium interactions held steady of the period; dosage failure rates dropped in all states, and duplicative therapy failure rates rose. The time trend for the latter criterion is shown in Exhibit 3.4 below. Since there is no clinical evidence to suggest that combination ACE inhibitor therapy improves patient outcomes, this trend warrants further investigation.

**Exhibit 3.4**

**HCFA DUR Demonstration Evaluation**
**DUR Criteria Failures Associated with ACE Inhibitors in the Four Study States, 1989-1996**
**(% persons failing criteria per year)**

| DUR Criterion/State | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 |
|---|---|---|---|---|---|---|---|---|
| Concurrent use of Multiple Ace Inhibitors | | | | | | | | |
| Georgia | | | | | | 4.29 | 3.59 | 6.53 |
| Iowa | 2.32 | 2.25 | 2.10 | 3.15 | 2.74 | 3.25 | 3.21 | 3.94 |
| Maryland | 3.20 | 3.39 | 3.19 | 3.70 | 3.33 | 3.98 | 4.11 | 4.57 |
| Washington | 2.66 | 3.00 | 2.60 | 2.93 | 2.99 | 3.25 | 4.98 | 5.26 |

*Source:* Medicaid drug data from Georgia, Iowa, Maryland, and Washington, 1989-1996.

*Antidepressants.* The nine DUR criteria for antidepressants are: adult maximum daily dose, concurrent use of multiple antidepressants or stimulants, and drug-drug interactions with antimuscarinic/antispasmodic agents, monoamine oxidase inhibitors, guanethidine/guanadrel, or clonidine. Specific antidepressants are contraindicated for use with antiparkinsonians and antiarrhythmics. Screen failure rates vary widely by criterion as discussed previously in reference to Exhibits 3.2 and 3.3. Failure rates are generally, but not exclusively, higher among the at-risk population subgroups. There are no consistent time trends in the person-level data: criteria associated with high (low) failure rates in 1989 tend also to be high (low) in 1996. However, as shown in Exhibit 3.5, when the data are aggregated at the claim level, two criteria show consistently increasing rates over the study period: concurrent use of antidepressants w/stimulants, and tricyclic antidepressants or maprotiline with clonidine. The largest increase in concurrent usage of antidepressants w/stimulants occurred in Iowa (a 33-fold increase over eight 7 years). The fact that person-based DUR failure rates for these two criteria held steady would imply significant growth in numbers of prescriptions per person and perhaps poorer therapeutic management as a result.

**Exhibit 3.5**

**HCFA DUR Demonstration Evaluation**
**DUR Criteria Failures Associated with Antidepressant Use in the Four Study States, 1989-1996**
**(% claims failing criteria per year)**

| DUR Criterion/State | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 |
|---|---|---|---|---|---|---|---|---|
| Concurrent use of Antidepressants w/stimulants | | | | | | | | |
| Georgia | | | | | | 0.13 | 0.31 | 0.76 |
| Iowa | 0.04 | 0.04 | 0.08 | 0.22 | 0.43 | 0.57 | 0.87 | 1.34 |
| Maryland | 0.09 | 0.07 | 0.10 | 0.15 | 0.26 | 0.43 | 0.69 | 1.21 |
| Washington | 0.10 | 0.09 | 0.12 | 0.21 | 0.32 | 0.36 | 0.43 | 0.57 |
| Tricyclic Antidepressants or Maprotiline with Clonidine | | | | | | | | |
| Georgia | | | | | | 3.51 | 3.60 | 4.24 |
| Iowa | 2.14 | 2.48 | 2.08 | 2.04 | 2.53 | 2.78 | 3.42 | 3.69 |
| Maryland | 3.56 | 3.73 | 3.80 | 3.49 | 3.22 | 4.17 | 4.95 | 5.20 |
| Washington | 2.31 | 2.29 | 2.86 | 3.28 | 3.49 | 4.20 | 4.98 | 5.64 |

*Source:* Medicaid drug data from Georgia, Iowa, Maryland, and Washington, 1989-1996.

*Antipsychotics.* There are seven DUR criteria in the antipsychotics category: adult and pediatric maximum daily doses, maximum duration, the concurrent use of more than one antipsychotic, and drug-drug interactions with antiparkinsonians, levodopa, or three or more antimuscarinic/ antispasmodics. Use of antipsychotics among very young persons is also contraindicated. Screen failure rates range from zero to 16 percent with significant differences across the states and among the risk groups within each state. The only consistent patterns are seen in the case of multiple antipsychotics and antipsychotic dosage. The DUR failure rates for both exceed 10 percent for all years in all four states. However, as shown in Exhibit 3.6 below, the trends differ markedly between the two, with rising failure rates for multiple antipsychotic use and declining rates for dosage failures. This inverse relationship may imply some level of substitution (high dosage monotherapies replaced

**Exhibit 3.6**

### HCFA DUR Demonstration Evaluation
### DUR Criteria Failures Associated with Antipsychotic Use in the Four Study States, 1989-1996
### (% persons failing criteria per year)

| DUR Criterion/State | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 |
|---|---|---|---|---|---|---|---|---|
| Antipsychotic Drugs, Dosage | | | | | | | | |
| Georgia | | | | | | 7.52 | 6.62 | 6.26 |
| Iowa | 12.3 | 12.6 | 13.0 | 12.6 | 9.46 | 8.40 | 8.06 | 7.46 |
| Maryland | 12.2 | 12.8 | 12.0 | 11.6 | 12.5 | 11.2 | 9.37 | 8.64 |
| Washington | 12.3 | 11.4 | 9.27 | 9.37 | 8.99 | 8.30 | 6.57 | 5.92 |
| Multiple Antipsychotics | | | | | | | | |
| Georgia | | | | | | 10.9 | 11.0 | 12.1 |
| Iowa | 12.9 | 12.7 | 12.4 | 12.5 | 11.7 | 14.7 | 14.8 | 16.3 |
| Maryland | 10.1 | 9.87 | 10.1 | 10.4 | 10.1 | 13.0 | 13.8 | 14.1 |
| Washington | 11.6 | 11.5 | 11.9 | 11.9 | 11.8 | 14.8 | 15.5 | 16.0 |

*Source:* Medicaid drug data from Georgia, Iowa, Maryland, and Washington, 1989-1996.

with lower dosage polytherapies). The rising rate of failures for the multiple agent criterion is unexpected since newer drugs such as clonazepine or risperidone are commonly replacing, rather than being added to, the older antipsychotic regimen.

***Benzodiazepines.*** The eight criteria in this class are: adult maximum daily dose, maximum duration, short-term use for insomnia, duplicative therapy, and concurrent use with other anxiolytic drugs or sedative/hypnotic drugs or antihistamines. Long half-life benzodiazepines are also contraindicated in the elderly or residents of long-term care facilities. Three of these criteria are highlighted in Exhibit 3.7 below. Of particular note are the low failure rates for these benzodiazepine screens in Georgia relative to the other states. This pattern is unique to benzodiazepines and is not found in other drug classes. Why? One plausible explanation is that Georgia has a very stringent prior approval process for benzodiazepines—just three prescriptions or refills per calendar year with mandatory prior authorization. Iowa implemented a prior authorization requirement for benzodiazepines in 1992, but the impact is minor compared to Georgia. Neither Washington nor Maryland had such targeted policies during the study period.

Also of interest in Exhibit 3.7 are the declining failure rates for all thee of benzodiazepine screens in Iowa, Maryland, and Washington. The declines are greatest in the case of the dosage screen. In Washington, for example, the DUR failure rate for benzodiazepine dosage fell from 11.5 to 3.8 percent. By contrast, both Iowa and Maryland saw much greater declines in multiple benzodiazepine screen failures than was the case in Washington. Despite the improvements in drug therapy implied by declining failure rates for certain benzodiazepines DUR criteria, this class of drugs remains notable as having the highest failure rates of any of the eight drug groups for all states in all years.

***Calcium Channel Blockers.*** There are nine DUR criteria for calcium channel blockers: adult maximum daily dose, a maximum duration for nimodipine, duplicative therapy, and drug-drug interactions with cyclosporine or beta-adrenergic blocking agents. Verapamil-drug interactions are

**Exhibit 3.7**

**HCFA DUR Demonstration Evaluation**
**DUR Criteria Failures Associated with Benzodiazepine Use in the Four Study States, 1989-1996**
**(% persons failing criteria per year)**

| DUR Criterion/State | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 |
|---|---|---|---|---|---|---|---|---|
| Benzodiazepines, Dosage | | | | | | | | |
| Georgia | 11.5 | | | | | 2.27 | 2.04 | 1.77 |
| Iowa | 6.42 | 10.4 | 9.44 | 8.97 | 6.41 | 4.89 | 4.08 | 3.82 |
| Maryland | 12.1 | 6.33 | 5.28 | 4.88 | 4.54 | 3.78 | 3.76 | 3.54 |
| Washington | | 8.91 | 7.12 | 7.61 | 7.66 | 6.35 | 5.81 | 5.34 |
| Estazolam, Flurazepam, Temazepam, Triazolam Duration | | | | | | | | |
| Georgia | | | | | | 15.6 | 20.9 | 26.3 |
| Iowa | 57.4 | 56.1 | 56.0 | 54.6 | 54.4 | 52.9 | 52.0 | 53.7 |
| Maryland | 49.8 | 48.0 | 47.8 | 43.8 | 41.8 | 41.8 | 44.5 | 44.1 |
| Washington | 45.5 | 44.0 | 41.9 | 39.4 | 36.1 | 35.0 | 36.5 | 38.7 |
| Multiple Benzodiazepines | | | | | | | | |
| Georgia | | | | | | 2.13 | 2.40 | 2.60 |
| Iowa | 16.7 | 15.8 | 15.4 | 15.0 | 14.7 | 12.3 | 11.4 | 11.4 |
| Maryland | 16.8 | 16.2 | 15.8 | 13.7 | 11.8 | 10.4 | 9.92 | 9.95 |
| Washington | 13.2 | 12.5 | 12.5 | 11.6 | 9.97 | 10.6 | 12.6 | 13.1 |

*Source:* Medicaid drug data from Georgia, Iowa, Maryland, and Washington, 1989-1996.

noted for digoxin and disopyramide or flecainide. Bepridil is approved only as second-line therapy for angina. Examination of the data in Appendix A provided little evidence of trends either up or down in failure rates for any of the calcium channel blocker DUR screens over the study period. Failure rates for four of the screening criteria in this therapeutic category were consistently below 1.0 percent in all states/years (adding verapamil to digoxin, initial therapy of bepridil, calcium channel blockers with cyclosporine, and verapamil with disopyramide or flecanide). One anomaly is the increasing rates of claims failing the calcium channel blocker with cyclosporine screen as shown in Exhibit 3.8 below. In Washington, the failure rates for this criterion increase from 0.35 percent of claims in1989 to 2.4 percent of claims by 1996. It should be noted that from a clinical standpoint this criterion is not an absolute contraindication, but care must be taken to adjust the dosages of both medications based on patients' blood levels. Although a serious interaction is possible, cyclosporine is usually prescribed only by transplant specialists, and such patients are carefully monitored.

***Digoxin.*** There are five digoxin DUR criteria: maximum daily maintenance dose, duplicative therapy, and drug-drug interactions with amiodarone/quinidine without a decrease in digoxin dose, high dose digoxin with amiodarone/quidine/propafenone, or cyclosporine. As with the calcium channel blocker criteria, screen failure rates for digoxin criteria are relatively patternless with fewer than four percent of users violating any screen in this therapeutic category except for adult and pediatric dosage (see

**Exhibit 3.8**

**HCFA DUR Demonstration Evaluation**
**DUR Criteria Failures Associated with Calcium Channel Blocker Use in the Four Study States, 1989-1996**
**(% claims failing criteria per year)**

| DUR Criterion/State | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 |
|---|---|---|---|---|---|---|---|---|
| Calcium Channel Blockers w/ Cyclosporine | | | | | | | | |
| Georgia | | | | | | 0.59 | 0.73 | 0.86 |
| Iowa | 0.92 | 1.31 | 1.44 | 1.66 | 2.06 | 2.29 | 2.14 | 2.27 |
| Maryland | 0.27 | 0.24 | 0.31 | 0.30 | 0.70 | 1.06 | 1.05 | 1.09 |
| Washington | 0.35 | 0.52 | 0.69 | 0.91 | 1.19 | 1.9 | 2.28 | 2.40 |

*Source:* Medicaid drug data from Georgia, Iowa, Maryland, and Washington, 1989-1996.

Exhibit 3.9 below). Of particular note here is Washington state where the prevalence of dosage failure rates exceeded one in four digoxin users in 1989. Although the failure rate subsequently dropped to below 10 percent of users by 1996, that was still more than double the rate found in any other state in that year. The other three states also recorded declines in digoxin dosage failures over the study period, but of more modest proportions.

**Exhibit 3.9**

**HCFA DUR Demonstration Evaluation**
**DUR Criteria Failures Associated with Digoxin Use in the Four Study States, 1989-1996**
**(% persons failing criteria per year)**

| DUR Criterion/State | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 |
|---|---|---|---|---|---|---|---|---|
| Digoxin: Adult and Pediatric Dosage | | | | | | | | |
| Georgia | | | | | | 3.15 | 3.08 | 2.66 |
| Iowa | 8.80 | 9.10 | 8.47 | 8.35 | 6.47 | 5.05 | 4.07 | 3.89 |
| Maryland | 10.8 | 11.2 | 8.92 | 7.48 | 6.38 | 4.87 | 4.93 | 4.90 |
| Washington | 27.4 | 23.0 | 18.1 | 17.6 | 17.4 | 14.7 | 12.8 | 9.26 |

*Source:* Medicaid drug data from Georgia, Iowa, Maryland, and Washington, 1989-1996.

***H2RAs.*** The eight criteria for H2RAs are: adult maximum daily dose, duration of acute therapy or GERD therapy, concurrent use of multiple H2Ras, or with misoprostol, omeprazole or sucralfate, and sucralfate therapy cannot be concurrently initiated. Drug-drug interactions are listed for ketoconaole, and it is inappropriate to add cimetidine to phenytoin, methylxanthines, warfarin or procainamide. The H2RAs present the second highest rates of DUR screen failures of all the drug classes following benzodiazepines. This therapeutic group also exhibits the most interstate differences. Some of the more dramatic patterns and trends are captured in Exhibit 3.10 on the next page. This table shows failure rate trends for three screens each at the person-level and the claims-level. All six are associated with generally declining failure rates, but the magnitude of the declines was much greater at the person level for the first three (H2RA and sucralfate on the same day, duration of therapy, and GERD duration of therapy), whereas the declines were greater at the claims level for the second set (drug-drug

**Exhibit 3.10**

### HCFA DUR Demonstration Evaluation DUR Criteria Failures Associated with H2RA Use in the Four Study States, 1989-1996 (% persons failing criteria per year)

| DUR Criterion/State (persons) | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 |
|---|---|---|---|---|---|---|---|---|
| H2RA and Sucralfate on the same day | | | | | | | | |
| Georgia | | | | | | 1.66 | 0.98 | 0.15 |
| Iowa | 5.98 | 5.6 | 5.39 | 3.99 | 2.56 | 2.25 | 1.70 | 1.31 |
| Maryland | 6.20 | 5.19 | 4.26 | 3.10 | 2.58 | 2.11 | 1.61 | 1.34 |
| Washington | 5.41 | 4.78 | 4.19 | 3.16 | 2.38 | 1.84 | 1.39 | 0.81 |
| H2RA: Acute Duration | | | | | | | | |
| Georgia | | | | | | 2.11 | 3.53 | 3.30 |
| Iowa | 13.2 | 12.9 | 12.2 | 11.7 | 7.07 | 5.57 | 5.89 | 5.51 |
| Maryland | 8.57 | 7.53 | 6.76 | 5.12 | 5.15 | 5.26 | 5.88 | 4.85 |
| Washington | 11.3 | 8.97 | 8.05 | 7.53 | 7.02 | 6.59 | 6.80 | 7.09 |
| H2RA: GERD Duration | | | | | | | | |
| Georgia | | | | | | 3.61 | 2.57 | 0.83 |
| Iowa | 10.5 | 11.4 | 11.9 | 12.2 | 3.48 | 2.60 | 3.68 | 4.12 |
| Maryland | 12.7 | 13.5 | 12.8 | 7.19 | 8.30 | 9.17 | 11.0 | 12.7 |
| Washington | 13.5 | 15.5 | 16.1 | 16.1 | 16.6 | 17.0 | 19.0 | 21.8 |

| DUR Criterion/State (claims) | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 |
|---|---|---|---|---|---|---|---|---|
| Cimetidine w/Phenytoin, Methylxanthines, Warfarin or Procainamide | | | | | | | | |
| Georgia | | | | | | 0.41 | 0.43 | 0.28 |
| Iowa | 0.85 | 0.80 | 0.74 | 0.69 | 0.62 | 0.62 | 0.55 | 0.41 |
| Maryland | 0.51 | 0.43 | 0.41 | 0.36 | 0.33 | 0.30 | 0.26 | 0.19 |
| Washington | 1.03 | 0.89 | 0.74 | 0.62 | 0.60 | 0.51 | 0.44 | 0.32 |
| H2RA w/Misoprostol, Omeprazole, Sucralfate | | | | | | | | |
| Georgia | | | | | | 6.71 | .86 | 3.56 |
| Iowa | 13.8 | 15.0 | 14.7 | 13.5 | 8.45 | 8.48 | 8.58 | 8.08 |
| Maryland | 13.7 | 13.7 | 13.1 | 12.2 | 11.5 | 10.7 | 10.1 | 8.35 |
| Washington | 14.2 | 13.3 | 13.6 | 12.9 | 11.8 | 11.0 | 10.0 | 7.80 |
| H2RA: Adult Dosage | | | | | | | | |
| Georgia | | | | | | 7.92 | 6.60 | 6.46 |
| Iowa | 33.8 | 34.8 | 35.4 | 35.4 | 4.2 | 11.9 | 13.2 | 13.2 |
| Maryland | 26.3 | 25.2 | 23.1 | 15.6 | 18.4 | 19.2 | 20.6 | 21.3 |
| Washington | 34.9 | 33.3 | 33.1 | 33.4 | 33.5 | 32.4 | 31.7 | 35.5 |

*Source:* Medicaid drug data from Georgia, Iowa, Maryland, and Washington, 1989-1996.

interactions between cimetidine and phenytoin, methylxanthines, warfarin, or procainamide, and between any H2RA and misoprostol, omeprazole or sucralfate). Possible reasons for these declines are the shift in treatment of gastric ulcers to a standard H. Pylori therapy, an increase in the use of omeprazole and other proton pump inhibitors, and the introduction of over-the-counter H2RA agents.

Under the GERD duration criterion, Washington is a clear anomaly with consistently increasing failure rates (from 13.5 to 21.8 percent of users between 1989 and 1996) whereas failure rates for this criterion declined or remained stable in the other states. Drug claims failing the H2RA adult dosage screen exhibit the greatest variation among the four states. In 1996, 35 percent of H2RA claims failed this criterion in Washington, again the highest of the four states (comparable rates were Maryland, 21 percent, Iowa, 13 percent, and Georgia 6 percent). In Iowa there was a sharp drop off in failure rates for several H2RA criteria in 1992, the same year that the state began requiring prior authorization for H2RAs.

*NSAIDs.* There are twelve criteria for non-steroidal anti-inflammatory drugs. These are: adult maximum daily dose, maximum duration for mefenamic acid/phenylbutazone,[7] duplicative therapy, and drug-drug interactions with high dosage methotrexate, select NSAIDs and oral anticoagulants, lithium, phenylbutazone or salicylates with oral andiabetics, triamterene and indomethacin, and salicylates or ketorolac with probenecid. Drugs in the NSAID therapeutic category are the most widely used among the eight drug groups examined in this study. Approximately half of all Medicaid recipients used one or more of these drugs per year from 1989 to 1996. Given the high prevalence of use, DUR failure rates are relatively low—less than one percent for most criteria. There appear to be very slight declines in failure rates for many of the NSAID criteria over the eight-year study period, but the pattern is not consistent across states. Exhibit 3.11 shows failure rates for two NSAID DUR criteria at the claims level. Here the varied interstate patterns are clearly pronounced. For example, the percent of claims failing the NSAID with oral anticoagulant interaction increased each year in every state, more than doubling (from 1.2 to 2.6 percent of claims) in Washington. In the case of another NSAID drug interaction criterion—salicylates with probenecid—failure rates increased steadily in Iowa (from 0.1 percent of claims in 1989 to 1.2 percent in 1996), but declined in the other three states. There are no ready explanations for such findings.

**Exhibit 3.11**

**HCFA DUR Demonstration Evaluation**
**DUR Criteria Failures Associated with NSAID Use in the Four Study States, 1989-1996**
**(% claims failing criteria per year)**

| DUR Criterion/State | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 |
|---|---|---|---|---|---|---|---|---|
| NSAIDs w/Oral Anticoagulants | | | | | | | | |
| Georgia | | | | | | 2.01 | 1.99 | 2.22 |
| Iowa | 2.23 | 2.39 | 2.88 | 3.01 | 3.15 | 3.31 | 3.15 | 3.47 |
| Maryland | 1.15 | 1.15 | 1.27 | 1.50 | 1.84 | 1.84 | 1.86 | 1.94 |
| Washington | 1.22 | 1.25 | 1.53 | 1.55 | 1.71 | 1.87 | 2.35 | 2.55 |
| Salicylates w/probenecid | | | | | | | | |
| Georgia | | | | | | 0.32 | 0.31 | 0.51 |
| Iowa | 0.14 | 0.52 | 0.30 | 0.74 | 0.72 | 0.69 | 1.13 | 1.18 |
| Maryland | 0.36 | 0.30 | 0.30 | 0.15 | 0.03 | 0.26 | 0.25 | 0.21 |
| Washington | 1.20 | 0.88 | 0.96 | 0.50 | 0.59 | 0.50 | 0.49 | 0.17 |

*Source:* Medicaid drug data from Georgia, Iowa, Maryland, and Washington, 1989-1996.

---

[7] Phenybutazone was withdrawn from the market in 1996.

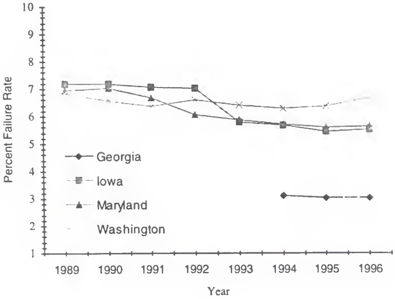
### 3.4.4 Summary Measures of DUR Failure Rates

Exhibits 3.12 to 3.15 show trends in DUR screen failures averaged across all 61 screening criteria. The average is computed by dividing the sum of recipients screened for all criteria by the sum of screen failures in each state/year.[8] Very clear patterns emerge from in these graphs.

Average DUR failure rates for the entire Medicaid population in each state are portrayed in Exhibit 3.12. The Iowa, Maryland, and Washington rates cluster between 5.5 percent and 7.2 percent of persons with one or more screen failures per year. Georgia's failure rates are half that, averaging just 3.1 percent per year. Both Iowa and Maryland exhibited significant improvements in DUR failure rates over the eight years of the study. In Iowa the rate fell from 7.2 percent to 5.5 percent, with most of the improvement registered in 1993, one year after the state adopted its prior authorization program for clozapine, H2RAs, sulcrafate, omeprazole, misoprostol, single source NSAIDs and benzodiazepines. The improvement in Maryland was more gradual with mean failure rates declining from 7.0 percent in 1989 to 5.7 percent in 1996. Washington ended 1996 about where it started in 1989 with DUR failure rates averaging 6.8 percent per year.

**Exhibit 3.12**

**HCFA DUR Demonstration Evaluation**
**Average Failure Rate for 61 DUR Screening Criteria in the Four Study States, 1989-1996**
**All Medicaid Recipients**



---

[8] Computing the average in this way weights each screening criterion by the number of persons affected. An alternative method of computing an all-screen average is to sum the failure rates for each screening criterion and then divide the result by 61. This later method gives each criterion equal weight regardless of the number of persons affected. The empirical results are similar either way. The equal-weight method produces higher failure rates by 1 to 2 percentage points, but the relative rankings of states and years are unaffected by choice of method.

The trends in DUR failures for multiple drug users (Exhibit 3.13) and adult diabetics (Exhibit 3.14) map closely those for the Medicaid population as a whole, except that the rates are higher across the board by 1 to 2 percentage points. These results are not unexpected. Persons using drugs within several therapeutic categories are at increased risk for various DUR criteria violations, particularly drug-drug interactions. Although only one DUR screen among the 61 specifically targets diabetes medications (phenylbutazone w/oral antidiabetic), the complications of this disease mean that adult diabetics typically take various medications which also increases their risk for DUR failures.

Exhibit 3.15 displays summary data on screen failure rates for the community-dwelling elderly subpopulation in each state. The findings shown here are perhaps the most dramatic of any presented thus far. Here too, Georgia led all states with the elderly failing DUR screens at an average rate of between 2.2 and 2.4 percent annually (slightly higher than for the entire Medicaid population). The real story is in the other three states where DUR failure rates for the elderly fell precipitously over the eight-year time span. In Iowa, the failure rate for elderly recipients in 1989 was equivalent to the

**Exhibit 3.13**

**HCFA DUR Demonstration Evaluation**
**Average Failure Rate for 61 DUR Screening Criteria in the Four Study States, 1989-1996**
**Multiple Drug Users**

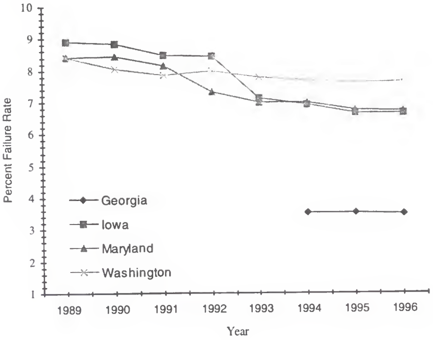
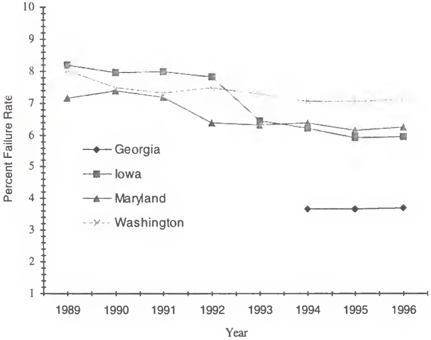


**Exhibit 3.14**

**HCFA DUR Demonstration Evaluation**
**Average Failure Rate for 61 DUR Screening Criteria in the Four Study States, 1989-1996**
**Adult Diabetics**



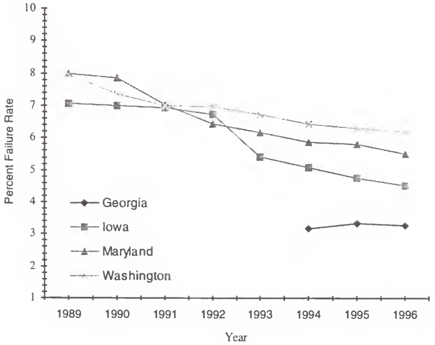
Medicaid population as a whole (7.1 versus 7.2 percent, respectively). But by 1996, the mean failure rate for the elderly was only 4.5 percent, a full percentage point lower than the population average. In Maryland and Washington, 1989 failure rates for the elderly were much higher than the population averages (8.0 verses 7.0 percent in Maryland, 8.0 versus 6.9 in Washington). By 1996, however, the rates for the elderly were below the population levels (5.5 versus 5.7 percent in Maryland, 6.2 versus 6.7 in Washington). This is a considerable achievement. Although there are probably many contributing factors involved here, the considerable attention devoted to drug problems and aging in both the scientific and lay press since the late 1980's must surely have played a major role.

## 3.5 Summary and Conclusions

This chapter has presented the findings from screening eight years of drug claims for Medicaid programs in Georgia, Maryland, Iowa, and Washington using a consistent DUR screener system developed by the Pennsylvania State University. The findings, presented in Appendix A and in Exhibits 3.1 to 3.15, represent a rich source of data for policy analysis of drug-related problems in these programs.

Exhibit 3.15

### HCFA DUR Demonstration Evaluation
### Average Failure Rate for 61 DUR Screening Criteria in the Four Study States, 1989-1996
### Community-Dwelling Elderly



Here is a summary of the most significant findings:

- Georgia had the lowest and Iowa the highest rates of drug utilization within the eight categories reviewed. In some drug classes usage rates were 50 percent higher in Iowa than Georgia.

- Drug use rose in all states over the study period. The largest increases were recorded for antipsychotics, ACE inhibitors, and calcium channel blockers.

- Georgia had the lowest mean failure rate at 3.1 percent of users in 1996 compared to rates of 5.4 to 6.7 percent in the other states that year.

- Significant improvement in DUR failure rates was recorded in Iowa (from a mean of 7.2 percent of persons failing DUR screens in 1989 to 5.5 percent failures in 1996) and Maryland (from a failure rate of 7.0 percent to 5.7 percent).

- Much of the improved performance is attributable to declining DUR failures among community-dwelling elderly. Over the eight years of the study, average DUR failure rates in this population declined by 35 percent in Iowa, 31 percent in Maryland, and 23 percent in Washington.

- Multiple drug users and adult diabetics had DUR failure rates 1 to 2 percentage points higher than the Medicaid population as a whole. Both groups showed improvement over the study period, but not as dramatic as for the community-dwelling elderly.
- DUR failure rates vary significantly by drug group and DUR criteria. Ten of the 61 criteria were associated with consistently high failure rates exceeding 8.0 percentage of persons and claims in every state/year reviewed. Thirteen were associated with consistently low failure rates below 0.25 percent of persons or claims per year.
- The drug groups with the highest failure rates cluster in the benzodiazepine, H2RA and NSAID classes. DUR criteria relating to dosage and duration of therapy problems are also associated with high failure rates.

The differentials in DUR failure rates highlighted here raise important policy questions. Why did failure rates fall substantially in Iowa and Maryland but not Washington (except for the elderly)? What specific factors explain the dramatic reductions in DUR failures among community-dwelling elderly? Might not these same factors be brought to bear to improve the drug use profiles of other high-risk groups? Why are the experts' opinion regarding appropriate dose and duration criteria followed much less often than their judgements regarding drug interaction and duplicative therapy criteria? Perhaps selected screening parameters should be empirically derived to better reflect actual prescribing practices. But if so, which ones?

Data alone cannot answer these questions. Descriptive studies such as this one are better at generating hypotheses than at testing them. It remains for future work to design research protocols in conjunctions with DUR databases to take full advantage of this valuable information source.

# 4.0 Paying for Pharmacists' Cognitive Services: The Project C.A.R.E. Demonstration

## 4.1 Background[1]

Pharmacists play a critical role in ensuring that patients receive and use medications appropriately. In addition to product selection, packaging, labeling, and basic patient counseling, pharmacists may interact with patients and/or prescribers to uncover a potential or existing prescribing problem, or to suggest a more effective, less toxic, or less costly drug regimen [Christensen et.al. (1996)]. As the link between the prescriber and the patient, the pharmacist is in a unique position to provide these types of enhanced, *cognitive* services.

Cognitive services provided by pharmacists can prevent a range of prescribing problems, including drug-drug interactions, drug-disease interactions, dosage errors, duplicative therapy, and patient noncompliance. The vulnerable populations served by Medicaid, which include indigent elderly people and children, are at higher risk for these prescribing problems and their associated adverse outcomes [Christensen et.al. (1996)]. These groups therefore are the most likely to benefit from an effective cognitive services program.

Researchers have found that pharmacists in a variety of clinical settings do provide some level of cognitive services, and there is evidence that cognitive services are effective in increasing appropriate prescribing and reducing costs. Most of these research results, however, were based on studies without randomized, controlled study designs. Despite the lack of rigorous studies in this area, many policy makers and professional organizations already have become supporters of cognitive services and, in general, a more proactive, patient-oriented role for pharmacists [Christensen et. al. (1996)]. Since 1974, Washington law, in fact, has required pharmacist counseling for all patients, regardless of payer.[2]

The traditional system of pharmacist payment provides no financial incentives for cognitive services; pharmacists are paid to dispense prescriptions, and activities such as cognitive services reduce the amount of time available for dispensing. The Cognitive Activities and Reimbursement Effectiveness (C.A.R.E.) demonstration project is based on the idea that if pharmacists are compensated for the time they spend providing cognitive services, the amount of cognitive services will increase. Increased amounts of cognitive services potentially may result in more appropriate, less costly prescribing and fewer drug-related, adverse health outcomes.

Project C.A.R.E. was initiated in October 1992, when the Health Care Financing Administration (HCFA) signed a cooperative agreement with the Washington Department of Social and Human Services (DSHS) and the University of Washington School of Pharmacy to implement a demonstration of payment for pharmacists' cognitive services. Project C.A.R.E. began reimbursing pharmacists for cognitive services (CS) provided to Medicaid beneficiaries in February 1994, with plans to close down the demonstration in January, 1995. With additional funding, CS payments were extended to September 1995.[3] The primary goals of Project C.A.R.E were the following [Christensen et. al. (1996)]:

---

[1] This section is based on site visits conducted in Washington in November, 1995, TAG reports, and the final report from Project C.A.R.E..

[2] This law, however, did not clearly define what was meant by counseling.

- to design a system of pharmacist reimbursement for cognitive services provided to Medicaid beneficiaries
- to study the impact of cognitive services on: (1) the number and type of prescribing problems identified and resolved; and (2) the costs of drug therapy
- to determine what pharmacist and pharmacy characteristics are associated with cognitive services provision

The demonstration involved the random assignment of pharmacies that applied to Project C.A.R.E. to intervention and control groups (see Section 4.3 for a more detailed description of implementation). The intervention group (Group A) was required to document cognitive services provided to Medicaid beneficiaries. Group A pharmacies received a payment for each cognitive service provided ($4 for interventions lasting less than six minutes, $6 for longer interventions), in addition to a small monthly stipend ($40) for participation in the project. The control group (Group B) *did not* receive a payment per cognitive service rendered, but these pharmacies received the monthly stipend, and still were required to document cognitive services provided to Medicaid beneficiaries. A second, silent control group (Group C) received *no* payments and *did not* document cognitive services provided [Christensen et. al. (1996)].

Project C.A.R.E owes much in its design to a precursor, the Pharmacist Intervention Project, conducted by the Washington State Pharmacy Association in 1989 in 100 pharmacies. This project was, itself, inspired by two earlier efforts: the first was a five-state demonstration (of which Washington was one state) designed by Rupp (1991), and the second was a model CS payment system designed by Christensen et al. (1994) to reimburse pharmacists for routine, brief, and episodic encounters.

---

[3] After September 1995, the state agreed to continue paying Group A pharmacists to provide CS as long as funds allocated to this item remained. However, all demonstration pharmacists stopped submitting the demonstration's CS documentation forms (although documentation on submitted claims continued), so that it would not be inaccurate to say that the demonstration ended in September even though the intervention continued.

## 4.2 Pharmacists and Prescribers: Attitudes Toward DUR and the Demonstration

### 4.2.1 Background

In addition to analyses which measured the economic and health impact of Project C.A.R.E., the evaluator also engaged in process analyses to understand the context in which the demonstration operated. The process analyses drew on several primary data collection activities; focus groups with pharmacists and prescribers, a pharmacy (pharmacy owner)/pharmacist survey, and targeted interviews during site visits. The survey, conducted in 1995,[4] collected information on pharmacy and pharmacist characteristics, pharmacist attitudes toward and experiences with DUR in the context of OBRA 1990 regulations, and pharmacist perceptions of the demonstration project. Surveys were mailed to all pharmacies participating in the demonstration: Treatment Group A (flat reimbursement for documenting patient counseling and additional reimbursement for each cognitive service provided); and Treatment Group B (flat reimbursement for documentation of CS only). A silent control group, Group C (comparison), was also sampled.

The focus groups, also held in 1995, were designed to better understand pharmacists' and prescribers' (physicians and nurse practitioners) understanding and attitudes toward DUR systems, attitudes toward and experiences with interacting with each other and with recipients, and suggestions for improvements in the system. Pharmacists and prescribers were recruited from the metropolitan Seattle area without regard to their exposure to the demonstration, so these groups do not directly address the demonstration and do not purport to be representative of attitudes across the state. They do, however, present interesting insights and recommendations that were obtainable only in a group discussion format. Taken together, the survey and focus group findings shed some light on the implementation of drug utilization review policy at the ground level.

This section of the report will describe survey and focus group responses to OBRA 1990 DUR requirements, attitudes toward, practices regarding, and payment issues surrounding cognitive services, implications for the pharmacist-prescriber relationship, pharmacist attitudes toward work, and recommendations for improvements in the prescribing process.

### 4.2.2 OBRA 1990 Requirements

The DUR questions in the survey for pharmacy owners were primarily concerned with costs associated with complying with OBRA 1990 requirements as well as the cost of operating on-site prospective DUR and counseling. The survey found that these requirements did not impose a large financial burden on participating pharmacies, although about half of the pharmacies saw their prescription costs increase. In contrast, focus group participants believed that DUR-associated computer upgrades and new equipment were very costly and forced pharmacists to forego other improvements. Pharmacists in one group agreed that the main impact of DUR was that it induced pharmacies to upgrade their software and to buy printers.

---

[4] Evaluation of Drug Use Review Demonstration Projects: Annual Report (1996). Findings from the Washington Pharmacy Survey, Cambridge, MA: Abt Associates, Inc.

The focus groups also discussed the impact of OBRA 1990 on pharmacist practice. Pharmacists were generally aware of OBRA 1990 requirements and noted that DUR-related educational materials such as newsletters and prescription printouts for patient education were considered helpful with counseling and billing information at the point of sale. They noted, however, that printed material, if handed to the patient without explanation, increased patient anxiety about side effects because of the detailed, liability-oriented nature of the information. Pharmacists noted that physicians feel undermined when anxious patients confront them with information contained in printouts that the physician has not previously mentioned.

Most physicians were unaware of OBRA 1990, but both physicians and nurse practitioners noted that pharmacists were taking a more active role in the prescribing process and that this change in the pharmacist role affected their professional relationships in both positive and negative ways. Nurse practitioners in particular noted that there had been more calls from pharmacists since OBRA 1990 was instituted.

Survey respondents generally expressed favorable attitudes toward OBRA 1990 prospective DUR requirements for proactive screening of all prescriptions: 69 percent agreed or strongly agreed that conducting prospective DUR is a valuable use of their time, and 83 percent thought prospective DUR helped to avoid adverse patient effects. The majority (79 percent) of pharmacists agreed that prospective DUR did not interfere with the patient- pharmacist relationship, 65 percent agreed that prospective DUR did not interfere with the pharmacist-prescriber relationship, but in response to the relationship between the patient and the physician, only 58 percent agreed that it did not interfere in patient-physician relationships. Two-thirds (66 percent) of pharmacists agreed that ProDUR is helpful in their communication with patients, and 55 percent agreed that ProDUR helped them in their communications with prescribers.

When Washington pharmacists were asked about their ability to prospectively screen prescriptions on-line, it appeared that most pharmacists had that capacity. Computers that flag interactions, for example, were seen as useful, because technicians alerted to problems they may not be trained to pick up on their own can then inform the pharmacist. Computers also provide drug history (depending on the inter-pharmacy network one has on-line), making it easier to pick up interactions, early refills, and drug seeking.

### 4.2.3 Attitudes Toward Cognitive Services

Survey respondents were asked specifically about their attitudes toward cognitive services (CS)[5] since many had participated in the demonstration. Attitudes toward cognitive services among survey respondents were generally favorable; 77 percent of pharmacists felt that CS helped facilitate communication with patients and 68 percent affirmed that CS did so with prescribers. Most of the survey participants (89 percent) responded that CS helps avoid serious adverse patient effects, and 80 percent responded that CS activities were supported by their managers. In comparison with their attitudes toward prospective DUR, more pharmacists agreed than disagreed that the provision of CS has increased the amount of time in face-to-face communication with patients (62 percent).

---

[5] The survey provided several examples of CS from OBRA 1990: instances when a pharmacist consults a prescriber (or other pharmacist); patient or care giver counseling; monitoring or educating patients; consulting external sources of prescription drug information.

Pharmacists in focus groups were asked to discuss counseling as an OBRA 1990 requirement. Their expressed attitudes toward counseling requirements were neutral. They noted that the counseling requirements under OBRA 1990 were no different than the state regulations under which they had been operating since the 1970s. Two pharmacist comments are illustrative: "We didn't have to have OBRA '90 to tell us to do these things;" and "OBRA '90 doesn't change my practice although when it's a Medicaid patient I counsel a bit more." They did acknowledge, however, that there is now more monitoring from the state to assure compliance with the law. One pharmacist noted that because the state prefers that pharmacists, not pharmacy technicians, perform counseling, pharmacists now delegate simple dispensing to technicians in an attempt to carve out time for the pharmacist to perform counseling.

Pharmacists participating in focus groups universally agreed that patient counseling is an important pharmacist role. As defined by focus group participants, its primary goals are to ensure patient compliance, to oversee the drug regimen when multiple providers are involved, and to field questions that the patient feels too uncomfortable to ask the doctor. Pharmacists stated that medications are much more complex than in the past and that they often need to reinforce complicated administration procedures (e.g. for injectables), making counseling particularly important. Pharmacists view themselves as more receptive to patient questions than doctors and more knowledgeable, at times, with regard to the whole prescribing picture: " I find senior citizens don't ask a lot of questions. They feel that the doctor's too busy...No one's looking out for them." Another pharmacist asserted: "Sometimes there are many doctors involved, too, and so you need to be the person watching out for them. You need to be the watch dog, because this doctor doesn't know what the hell the other is doing." On the issue of patient compliance, one pharmacist stated that the counseling session "is the last chance for us to double-check to make sure the patient is receiving the proper medication."

### 4.2.4 Payment for Provision of Cognitive Services

Pharmacy owners whose pharmacies were in Treatment Groups A and B of the demonstration were asked about their attitudes toward payment for provision of counseling. Less than half (39 percent) of respondents who were being paid the $40 per month by the demonstration for documenting cognitive services thought this was adequate reimbursement. In terms of reimbursement for each encounter, 66 percent of Group A were satisfied with the $4 reimbursement for a brief encounter, while only 37 percent of Group A thought that $6 was adequate or very adequate for an extended encounter (over six minutes).

Over 40 percent of staff pharmacists responded that they could not provide as much counseling as was needed. Eighty-six percent cited lack of time, four percent responded that the patient would not cooperate, and three percent cited lack of privacy. Interestingly, only seven percent of staff pharmacists said that they could not provide as much counseling as was necessary because payment was too low; the payment issue was apparently perceived as more problematic to pharmacy owners than to their staff. Over one quarter (27 percent) of staff pharmacists thought that the required documentation was burdensome.

Focus group participants discussed reimbursement for counseling in the context of the changing health care market. As one pharmacist put it: "The quality of counseling is going down the tubes because the practices of pharmacies are changing." Managed care, increased competition between pharmacies, and low vending reimbursement per prescription were cited as major contributors to the decline in high quality counseling. "It doesn't give enough time to the pharmacist to spend five or ten minutes quality time with that customer, to get involved with what they take [and] to call the physician." Moreover, new pharmacists are increasingly working alone, with no mentoring from more experienced pharmacists on the finer points of counseling. OBRA 1990, they recommended, should have built in reimbursement for the increased physician contact and counseling the legislation requires. Such reimbursement would ease the financial burden on pharmacists to dispense as many prescriptions as possible in the shortest period of time.

### 4.2.5 The Pharmacist-Prescriber Relationship

Demonstration project participants, who volunteered to be randomized (Groups A and B) had more favorable attitudes toward prescribers than did Group C pharmacists. A significantly higher number of pharmacists from Group A (80 percent) vs. Group C (68 percent), and somewhat higher than Group B (71 percent) agreed or strongly agreed that CS does not interfere with the pharmacist-prescriber relationship.

Perhaps the most significant finding from the Washington focus groups had to do with the changing pharmacist-prescriber relationship as a result of legislation requiring a more proactive pharmacist role in the prescribing process. On the whole, pharmacists and prescribers agreed that the benefits of increased communication between the two professional groups on behalf of their patients was beneficial to quality of care, but that tensions arose nonetheless, less so with nurse practitioners, and more so with physicians.

Nurse practitioners and pharmacists expressed mutual respect for one another. Nurse practitioners stated that they appreciated the phone calls, perhaps because they, like pharmacists, are trained to do counseling whereas doctors are trained primarily as diagnosticians. Pharmacists seemed to agree with one another that younger doctors tended to be more receptive to pharmacist calls than older ones. Physician participants, when asked if they regularly and rapidly responded to pharmacist calls, felt that they were, in fact, accessible.

When a prescribing problem is identified in the pharmacy, how the pharmacist handles the situation has an impact on physician opinions of their counseling work. Some pharmacists alert their patients prior to calling the doctor, while others call the physician prior to alerting the patient. Physicians in focus groups strongly asserted that they prefer the latter approach. To physicians, pharmacists are most helpful when they detect real prescribing problems and "annoying" or "interfering" when they call about non-indicated, or off-label prescribing, which are not uncommon practices. Physicians noted that pharmacists' knowledge of formularies is helpful, but that their "cookbook" approach to prescribing makes them jump to conclusions unnecessarily. Physicians feel that calling a potential prescribing problem to the attention of the patient without consulting doctors as to the drug's intended use undermines their authority and oversteps the boundaries within which the pharmacist operates. Pharmacists did acknowledge that some of their colleagues become overzealous, citing situations in which pharmacists warn pregnant women against taking drugs that the doctor feels are necessary to preserving the health of the woman. Each situation, pharmacists stated, should be assessed carefully to ensure that a physician's judgment is not questioned when there is a legitimate reason for the prescription.

### 4.2.6 Pharmacist Attitudes Toward Work

The survey found that participation in a demonstration with reimbursement for documentation an provision of CS had a favorable effect on Group A pharmacies. A significantly higher number of Group A pharmacists (80 percent) vs. Group C (61 percent) and somewhat higher than Group B (71 percent) agreed or strongly agreed that patient care would suffer if they failed to provide drug information to the physician. It may be that being part of a demonstration sensitized pharmacists to the importance of informing physicians of adverse outcomes; it may also be that those pharmacists choosing to participate in the demonstration came into it favorably disposed toward counseling.

Certainly, pharmacists participating in the focus groups supported patient counseling as an important pharmacist role with implications for the health of the patient and for the quality of prescribing. Pharmacists expressed regret that health care professionals in general are less likely to know their patients than in the past, making detection of drug problems more difficult, underscoring the need for consistent, high quality patient counseling.

### 4.2.7 Recommendations for Improved Prescribing

Physicians felt that pharmacists can be the most helpful to them when they offer lower cost drug options and assist with narcotics monitoring, especially when they are aware of a patient's drug seeking behavior. Overall, both physicians and pharmacists felt that more information was necessary to improve communication between the groups and ultimately, to improve prescribing. Pharmacists recommended that access to medical records, or at least diagnoses, would be helpful. Full computerization of patient drug history would also help, with doctors required to list all medications a patient is on. Physicians recommended that an information system be available to them and pharmacists alike that would identify drug interactions and list all patient medications, particularly for older patients who may be seeing several physicians and whose recall may not be as sharp. Pharmacists also recommended that they be reimbursed for the provision of counseling services.

### 4.2.8 Conclusion

Washington did not mount a demonstration that compared OPDUR and cognitive services, nor has Washington imposed an OPDUR system on its pharmacies serving Medicaid beneficiaries, so it is impossible to say whether pharmacists prefer one method of DUR over another. What can be said is that pharmacists responding to the survey and participating in focus groups felt that patient counseling was an important role that did not interfere in their relationships with physicians and other prescribers. It is acknowledged by both pharmacists and physicians that the more proactive role of pharmacists is not without its conflicts, particularly when pharmacists make determinations or provide information to beneficiaries that physicians feel encroaches on their turf. There is no doubt that both groups agree that more communication between them is beneficial to patients, and that more information on drug history would improve prescribing.

Pharmacists were clear that in a more competitive, less lucrative, more complex health care market, reimbursement for patient counseling is absolutely necessary. They also noted that the current atmosphere will make counseling more difficult to do unless it is reimbursed adequately.

## 4.3 Project C.A.R.E.: Implementation

As is usually the case in such demonstrations, investigators from the University of Washington confronted unforseen challenges in implementing Project C.A.R.E. Decisions on how to address these challenges balanced the need to assure the continued engagement of participants against the need to protect the integrity of the research design.

This section of the report focuses on the implementation of Project C.A.R.E., and includes a discussion of recruitment, enrollment, and interactions with pharmacies, the cognitive services payments, data management, staffing and the evaluation. Despite some implementation problems, for the most part, the project and evaluation design remained intact. Even so, when designing and interpreting analyses of CS payment effects, HCFA's evaluator and Project C.A.R.E. investigators had to account for the nature and timing of some such mid-course changes in the demonstration.

### 4.3.1 Recruitment, Orientation and Training

Project C.A.R.E. recruited pharmacy participants through direct mailings, advertisements in industry and state publications, and presentations to corporate officers of chain drug stores. To be eligible, pharmacies must have been serving primarily ambulatory patients and dispensing at least 50 Medicaid prescriptions per month. Pharmacies that were part of a staff-model HMO were not eligible for the study [Christensen et. al. (1996)].

Pharmacies that enrolled (the enrollment process is described in the next section) received training manuals and quick reference training guides, had access to a toll-free help number, and were invited to attend training sessions held around the state [Christensen et. al. (1996)]. At the training sessions (which were evaluated very favorably by participants), cognitive service interventions and documentation methods were discussed in the context of case studies. Additional, individualized training sessions and review were provided for several sites.

Project C.A.R.E. also issued a newsletter, *C.A.R.E. Talk*, to maintain contact with participating pharmacists. The Project Manager reported that the newsletter had been well-received, but that difficulties experienced with the toll-free line contributed to a certain amount of irritation among pharmacists. Calls during working hours were often handled by work study students who occasionally failed to channel questions in the correct direction, leading to some responses being delayed for several days. As a solution, the Project Manager took over many the hotline responsibilities herself.

### 4.3.2 Sampling and Enrollment of Pharmacies and Cohort Maintenance

Pharmacies were assigned to arms of the demonstration in geographic clusters, selected to minimize the frequency of prescriber cross-over (prescribers having prescriptions filled in both A and B pharmacies). Group C pharmacies were similarly clustered to minimize cross-over between C and A pharmacies.

In spite of substantial early interest in the demonstration (245 pharmacies volunteered to participate) only 160 met all enrollment criteria by February 1994, the first month in which participants were eligible to receive demonstration payments. Efforts to reach the original target of 200 Groups A and B pharmacies succeeded after two additional waves of enrollment, in April and September. By February 1995, only one pharmacy had withdrawn from the demonstration. When Project C.A.R.E. negotiated an extension through September 1995 of CS reimbursement (based solely on state funds), 90 percent of participating A and B pharmacies re-enrolled. Exhibit 4.3.1 below displays the sample sizes in Groups A and B during the three enrollment waves.

**Exhibit 4.3.1**

**HCFA DUR Demonstration Evaluation**
**Project C.A.R.E. Pharmacy Enrollment**

| Date | Total Enrolled | Group A | Group B |
|---|---|---|---|
| February 1994 | 160 | 86 | 74 |
| April 1994 | 193 | 107 | 86 |
| September 1994 | 200 | 110 | 90 |

*Sources:* Christensen (1996).

Staggered enrollment, unforeseen in Washington's initial design, offered a challenge to the evaluator for two reasons. First, although the majority of participating pharmacies were exposed to the demonstration for the "full" period (February 1994 through September 1995), 33 were exposed for 16 months and seven for 12 months. Analyses of demonstration effects had to adjust for phased entry by identifying each pharmacy with its enrollment cohort. Second, the learning curve for newly-enrolled pharmacies was bound to be shorter for later entrants. Project C.A.R.E. staff acquired knowledge in the first months of the demonstration that made them better able to address problems later on. A second wave of training, implemented in the fall, drew on this early experience.[6] The Project Manager noted that, by late summer of 1994, the project seemed to have "come into its own", meaning that staff had recognized and solved many of the start-up problems. This new level of confidence was seen to be at least as important as the second wave of training in increasing CS documentation flows. Thus later pharmacy cohorts should have reached a desired "steady state" in volume of CS and quality of CS documentation earlier than the first cohort, for which the first month or two of CS data may be of questionable value.

---

[6] Experience with training the first cohort of pharmacists convinced C.A.R.E. staff that the materials used were too "academic." In response, efforts were made to streamline and simplify training procedures and documents as much as possible. For example, the Project Manager replaced a complicated manual with a one-page laminated fact sheet describing the CS documentation process that pharmacists could keep near their computers for easy reference.

### 4.3.3 Implementing the CS Payment Intervention

Although a broad definition of cognitive services could include any professional actions a pharmacist might take except dispensing drugs, Project C.A.R.E. constrained the definition of CS eligible for payment under this demonstration. The C.A.R.E. investigators separated cognitive services into "routine pharmaceutical services" (services assumed covered by state Medicaid dispensing fees, such as clarifying incomplete or illegible prescriptions) and "value added pharmaceutical services" (VAPS). VAPS were themselves allocated to two categories: those related to dispensing (prospective drug use review, consultation with prescribers, patients and other pharmacists) and those not related to dispensing (academic detailing, in-service training, rendering patient advice on drug use when no prescription was involved). By choosing to pay only for VAPS related to dispensing, Project C.A.R.E. associated its intervention with point-of-sale prospective drug use review.

Project C.A.R.E. initially placed further restrictions on payment for CS:

- The Project required evidence for specific and identifiable potential drug therapy problems resolved during the course of routine dispensing activities and followed by changes in drug therapy (including a decision that no drug is to be dispensed).

- The Project would pay for patient counseling performed pursuant to an identified problem.

- Pharmacists were limited to payment for up to two documented cognitive service activities per patient per day.

Later, during discussions with HCFA, these restrictions (expecting the two CS per day limit) were relaxed to allow payment for all drug-related VAPS.

Project C.A.R.E. asked Groups A and B pharmacists to document all cognitive services, according to the problem identified, the intervention made and the result of the intervention. Project C.A.R.E. identified 24 problems, including 16 prescription-specific problems (dosage, duplication, interactions) and eight other problems. Seven of these were patient-specific (under-use, over-use, patient requires case management). Ten specific interventions (and one "other" intervention) were identified; three of these involved information gathering (literature reviews, chart reviews, lab tests), and seven involved interaction with prescribers, patients, other pharmacists or Medicaid. Finally, Project C.A.R.E. identified 12 results, including dispense as written, counsel or refer the patient and nine changes in a prescription (including do not dispense). Pharmacists were asked to identify only one "primary" problem, but they could identify one primary and more than one related intervention and result. In addition, pharmacists documented the time each CS took, and were asked to choose one of three levels of "morbidity risk" related to the CS, from "low," an "intervention likely to increase, at most, patient convenience," to "high," an "intervention likely to save the patient a subsequent ER visit or hospitalization."

### 4.3.4 Refinements and Modifications to the CS Payment Intervention

Project C.A.R.E. investigators retained the two-tier payment method for approved cognitive services and other features of the intervention throughout the demonstration. However, CS claims rejection rates were higher than expected early in the demonstration. Project C.A.R.E. investigators had provided DSHS with a list of valid CS codes, representing services that were eligible for payment (not part of dispensing) and that represented what were thought to be "logical" combinations of problem/ intervention/result (as listed in the CS documentation). In response to pharmacist concerns about claims rejections, Project C.A.R.E. investigators reviewed the logical and illogical CS codes and expanded the former and, in some instances, created new codes.

It is not clear that this expansion could have directly affected pharmacist behavior. There are hundreds of combinations of the three CS code components, and a working pharmacist would probably find it more rational to submit all CS for payment rather than spend time anticipating which would or would not be paid. Of course, the initial lag in payment for some of the rejected claims may have had a dampening effect on participation. The subsequent expansion of the list of eligible codes, accompanied by claims payments, might have partially reversed this effect, but the overall effects of this expansion on CS activity and documentation are likely to be unobservable.

Although not in itself a change in the intervention, Project C.A.R.E investigators noted an unexpectedly high frequency of the problem "case managed patient." Washington site visit interviewees observed that documented case management probably increased beginning in 1995 as a result of a DSHS decision to pay for up to two Mediset systems per year for high-risk groups. DSHS can restrict certain high-risk patients (e.g. patients who misuse drugs) to one pharmacy or one physician, who will presumably take more responsibility for risky cases and, by intervening early, reduce downstream hospitalization, thus saving costs overall. (Case managed patients referred to pharmacists are a small minority of all such patients enrolled in the Washington Medicaid program). Case managed patients are clearly at high risk of adverse outcomes linked to prescription drugs due to poor compliance. What is not known is whether or not the availability of CS payment, coupled with Medisets reimbursement, encouraged participating pharmacists to alter in any way the kinds of services provided to these patients after December 15, 1994.[7]

Experience with the two-tiered payment system has been mixed. On the one hand, Project C.A.R.E. investigators point to the fact that CS services average around six minutes in both A and B pharmacies to argue that the payment system recognizes a natural cut point in the time distribution of CS services. On the other hand, they concede that pharmacists facing lengthy and often complex interactions with prescribers, other pharmacists and the state Medicaid program, may well have been seriously underpaid.[8]

---

[7] One respondent recounted the experience of a rural Washington pharmacist who had encouraged a case managed patient to come each day to the pharmacy to receive his medications. After this patient missed two days, the pharmacist visited the patient's home, found the patient unconscious and succeeded in getting the patient hospitalized in time to prevent death. This story was meant to reinforce the respondent's view that pharmacists respond to the needs of case managed patients from both professional and human points of view, regardless of payment incentives.

[8] As a possible alternative approach that might more fairly compensate pharmacists for lengthy and complex interventions, C.A.R.E. investigators cited the new Wisconsin CS payment system which will pay according to a five-tiered set of rates.

### 4.3.5 Organization, Staffing and Data Management

Investigators from the University of Washington School of Pharmacy were expected to design, manage and evaluate Project C.A.R.E., under the direction of Principal Investigator Dale Christensen, Ph.D. Washington DSHS would process and pay CS claims submitted by participating pharmacies. In addition, 18 Area Coordinators (roughly one for every 10 participating pharmacies), pharmacists who had already agreed to participate in the demonstration, were to be recruited to intercede between their assigned pharmacies and Project C.A.R.E. staff, ensuring that pharmacies were submitting CS documents on a monthly basis and answering any CS documentation coding questions. Each Area Coordinator was to be paid $30 per assigned pharmacy for the period of the demonstration.

### 4.3.6 Organizational and Staffing Issues: Problems and Solutions

Competent staff in a demonstration can often overcome structural and procedural weaknesses in organization. There was considerable evidence that C.A.R.E. investigators learned quickly and addressed problems that emerged early in the demonstration.

Perhaps the main adjustment staff had to make in the demonstration's organizational structure was in replacing the functions performed by most of the Area Coordinators. Area Coordinators were participating pharmacists who had been recruited to provide help to pharmacists with CS documentation, claims submission and other day-to-day problems. As the Project Manager reported, some did a competent, thorough job but most performed unevenly. Pharmacists in the demonstration began routinely to circumvent their Area Coordinators, preferring to call C.A.R.E. management directly. Senior demonstration staff spent a considerable amount of time dealing directly with pharmacists' problems, particularly issues of delays in paying participation vouchers. It was ultimately decided that the function assigned to the area coordinators needed to be centralized and transferred to Project C.A.R.E. staff.

Project C.A.R.E. experienced turnover of key senior staff in two important areas: project management and data base design and management. The demonstration's original Project Manager left the position early during the implementation phase. In addition, the individual responsible for designing the Project C.A.R.E. data base left and moved out of the state early in the project. In both instances, steps were taken to maintain continuity and minimize disruption. A new project manager was recruited, and the original manager, whose skills and experience with CS documentation and training predated the demonstration, remained involved in pharmacist training. The data base designer maintained a 10 percent time commitment to Project C.A.R.E., albeit from a new position in Iowa. Fortunately, the project was successful in maintaining a core staff of analysts, including the Principal Investigator and Co-Principal Investigator, with experience dating from the demonstration's original conception.

### 4.3.7 Data Base Design and Implementation: Problems and Solutions

Some decisions made in planning data flows for the demonstration may have had negative consequences for the ability of both the University and the Abt team to evaluate Project C.A.R.E. Issues identified in early technical assistance site visits included:

- An apparent lack of attention during training to the importance of recording drug code and quantity information;

- Decisions to allow three different CS documentation methods: problems in the software to support computerized documentation are noted below — the main evaluation issue is the pharmacy-level effects of using these three different methods;

- Although tested in-house and in a few pharmacies, the automated CS documentation systems were not adequately field tested — after training, pharmacists were expected to utilize documentation procedures;

- A decision, based on confidentiality concerns, not to incorporate extra data elements (e.g. patient IDs) in the CS documentation that would facilitate linkages with CS service documents and prescription claims — Washington planned to depend on other data elements (prescription codes and dates, for example) to accomplish linkages.

By providing a choice of three CS documentation methods (the "Stotler" and "Lloyd" software systems, and paper documentation), the project, while trying to accommodate different pharmacists' preferences and capabilities, generated problems that diverted staff energies early in the operational phase of the demonstration. Staff had expressed concern at low rates of CS documentation in some pharmacies. However the link between this problem and the Stotler program was not discovered until a pharmacist submitted a Stotler disk with a note that the disk should contain eight documents. Staff found only one document. In following up, staff discovered that nearly 2,000 CS documents had not been recorded because pharmacists were failing to utilize a particular sequence of key strokes when entering the data. Stotler agreed to travel to the pharmacies that used his system to provide retraining. Pharmacies were also given the option of switching to paper documentation. Project C.A.R.E. staff believe that their audit, conducted after the problem was detected, recovered virtually all of the missing CS documents.

### 4.3.8 Data Flows

Initially, the demonstration was designed to generate three separate data flows:

1. CS service claims, submitted for payment by Group A pharmacies directly to DSHS; claims formats matched DSHS prescription drug claims, and it was assumed that most billing would be done electronically.

2. Project C.A.R.E. participation vouchers, submitted for payment by Groups A and B pharmacies directly to DSHS

3. CS documentation forms, submitted by Groups A and B pharmacies to the University of Washington.

Pharmacists assigned to Groups A and B had a choice of methods for documenting CS. They could elect to complete forms manually, or they could submit diskettes. Project C.A.R.E. provided two software options to participating pharmacies: the Stotler program and the Lloyd program.[9]

The Project C.A.R.E. core data base was designed to track CS documents as they moved through the system (C.A.R.E. suspended records until corrections were made, internally or externally by pharmacists) and to store the corrected records. Other reports and documents more fully describe this data base.

### 4.3.9 Program Evaluation

The University of Washington proposed to evaluate the effects of paying for CS on:

- numbers and types of drug-related problems identified, corrected or resolved;
- the costs of drug therapy; and
- the number and types of drug-related problems identified through RDUR screens.

The University also proposed to assess:

- factors associated with pharmacists' decisions regarding participation in Project C.A.R.E.;
- pharmacist, prescriber and patient characteristics associated with prescriptions that require CS; and
- the computerized methods of documenting CS implemented for the demonstration.

### 4.3.10 Conclusion

Project C.A.R.E. succeeded in fielding a demonstration that paid for CS and convinced pharmacists to document CS, with minimum attrition of participating pharmacies and probably no more than the usual amount of dissatisfaction with logistical problems, mostly in lags in paying participation vouchers during the first months of the demonstration. Some problems of implementation could perhaps have been anticipated and addressed earlier; for example, more intensive training and a shake-down period of a month or two might have detected problems in how pharmacists used the Stotler program. Others, such as the need to re-evaluate and expand the reimbursable CS codes, would have been difficult to forecast.

In their final report, Project C.A.R.E. investigators [Christensen et al., 1996] found that paying for CS did increase the volume of documented CS. However, much recorded CS activity involved patient counseling and education, rather than interactions with prescribers and other pharmacists.

---

[9] Earlier, the evaluation team identified variation among pharmacies in how CS would be documented as a potential influence on the completeness and content of the documentation, and therefore an issue to be addressed in the evaluation.

- In Project C.A.R.E. pharmacies, CS rates (per dispensed prescription) rose over time.

- Explicit documentation policy, lower workload, supportive relationships between pharmacists, patients and prescribers correlated with more CS.

- Over an 18-month period, Group A pharmacies documented from 1.3 to 2.4 CS per 100 dispensed prescriptions, compared to 0.7 to 1.0 CS in Group B. Per pharmacy rates of CS differed in a similar manner, averaging 136 for Group A and 58 for Group B.)

- However Group B pharmacists (43.7 percent) were more likely to consult a prescriber than Group A pharmacists (31.4 percent).

- About half of all CS provided by Groups A and B pharmacists were for patient-related problems, one-third for drug-related problems, 18 percent were for prescription-related problems.

- Drug therapy changes occurred in 27 percent of all CS. These changes were almost always accompanied by evidence of communication with prescriber. Average cost savings for change was estimated to be $13.05 (this includes dispensing and CS payments, reflects Medicaid savings before rebates).

- Project C.A.R.E. investigators estimated savings by type of CS: drug/drug regimen change ($17.70), discontinued drug ($36.88), do not dispense ($40.70). When the action involved adding drug therapy , the added costs averaged ($71.32). Average cost saving associated with CS that resulted in a change: $13.05 per CS , with averages of $11.45 in Group A and $15.33 in Group B.

- Average CS time in all participating pharmacies was 7.5 minutes, with fewer than 6 percent requiring 20 minutes or more. Group A time exceeded Group B (7.9 minutes, compared to 6.5 minutes)

In general terms, the demonstration has shown that CS payment is probably replicable in other states, although lessons learned here might help suggest improvements to the model. An expansion of the payment system to more fairly compensate lengthy and complex interventions has been mentioned. Although Washington's three-dimensional definition of what CS is offers a useful framework, the experience gained in Project C.A.R.E. does not define the appropriate level and content of CS documentation. For this demonstration, the purpose of documentation was to generate data for the evaluation.

Ultimately, of course, the success of this cognitive services model depends on its costs and its impact on behavior and health-related outcomes. Analyses completed for this evaluation report will help the reader come to some conclusions about the worth and viability of such a program if it were implemented in other states.

## 4.4 Findings on the Impact of the Cognitive Services Demonstration on Downstream Utilization and Costs

### 4.4.1 Background

Medicaid pays a substantial amount for beneficiaries' prescription drugs. In 1995, the average drug expenditure per Medicaid recipient was $270, leading to a total expenditure of $9.8 billion, an 120 percent increase from 1990.[10] For these financial reasons alone, Medicaid drug expenditures deserve active management and control.

However, there is a second compelling reason for more active review of Medicaid drug transactions: the protection of patients and improvement in the quality of drug therapy patients receive. Less-than-optimal drug utilization is a significant cause of health problems. By one estimate, 28.2 percent of all hospital admissions in the United States are associated with drug-related morbidity and mortality.[11]

This study evaluates Project C.A.R.E., the Washington demonstration of payment for cognitive services which took place between February 1994 and September 1995. In theory, payment for cognitive services should increase pharmacists' efforts to ensure optimal drug utilization, thereby improving drug prescribing and patient compliance, and thereby decreasing Medicaid recipients' health services utilization (specifically utilization involving diagnoses related to suboptimal drug use). This study tested the hypothesis that payment for cognitive services caused a decrease in non-drug health services utilization among individuals aged 65 and over who were dually eligible for both the Medicare and Medicaid programs. It did not find evidence in support of this hypothesis.

Previous research has indicated that between .8 percent and 2.6 percent of new prescriptions have a problem requiring pharmacist intervention and that the cost of managing drug-related morbidity and mortality is as high as $76.6 billion annually.[12,13] Some believe that increased pharmacists' services may ameliorate these problems.

Hepler and Strand define pharmacists' services as "the responsible provision of drug therapy for the purpose of achieving definite outcomes that improve a patient's quality of life" [Hepler & Strand (1990)]. These services, which are focused on the patient, both involve monitoring prescriptions and educating patients.

---

[10] U.S. Bureau of the Census, *Statistical Abstract of the United States: 1997*, Washington, DC, 1997.

[11] Johnson, Jeffrey A. and Bootman, J. Lyle. "Drug-Related Morbidity and Mortality: A Cost of Illness Model," *Arch Intern Med*; 1995; 155:1949-56.

[12] For percent of prescriptions with problems requiring pharmacist intervention see: Rupp, Michael T. "Value of Community Pharmacists' Interventions to Correct Prescribing Errors," *The Annals of Pharmacotherapy*; 1992; 26:1580-84; Rupp, Michael T. "Evaluation of Prescribing Errors and Pharmacist Interventions in Community Practice: An Estimate of 'Value Added'," *American Pharmacy*; 1988; NS28(12):766-770; and Dobie III, Richard L. And Rascati, Karen L. "Documenting the Value of Pharmacist Interventions," *American Pharmacy*; 1994; NS34(5):50-54.

[13] For cost of drug-related morbidity and mortality, see Johnson, Jeffrey A. And Bootman, J. Lyle. "Drug-Related Morbidity and Mortality: A Cost of Illness Model," *Arch Intern Med*; 1995; 155:1949-56.

There is some quantitative evidence that pharmacists' services can improve patient care. Jameson, VanNoord, and Vanderwoud found that an intensive program in which pharmacists met with both patients and doctors reduced drug costs, simplified drug regimens, lowered rates of side effects, and improved patient compliance and understanding [Jameson et al. (1995)]. There is also considerable evidence that pharmacists' services save money. Estimates of the gross value of pharmacists' interventions range from $2.32 to $7.15 per prescription [Rupp (1992, 1988, 1994)].

Some argue that pharmacists' compensation structure should be altered so that pharmacists are paid not only for dispensing medication but also for the cognitive services they provide. According to survey results collected and analyzed by Miller and Ortmeier, pharmacists report that the greatest barrier to their providing pharmacists' services was the fact that they were compensated for dispensing prescriptions and not for providing such services [Miller et al. (1995)].

Project C.A.R.E. offered an opportunity both to test the implementation of payment for CS and to test whether payment for CS did indeed translate into measurable increases in pharmacists' services, decreases in drug utilization and prescription errors (as measured by the PSU screener), and decreases in downstream health services utilization, the subject of the current study.

### 4.4.2 Data and Methods

This study investigated whether Medicaid recipients who filled their prescriptions at treatment pharmacies had lower health services utilization than recipients who filled their prescriptions at control pharmacies. For such an effect to exist, two things must have happened. First, the existence of per-service payments must have led pharmacists to perform more cognitive services; even without payment, pharmacists make efforts to guarantee that Medicaid recipients receive safe and effective drug therapy.[14] Second, these additional services must have measurably affected utilization. Either they must have substituted directly for other utilization or they must have been additional services with a protective effect. Either by changing prescriptions or by changing recipients' behavior, cognitive services must have improved subsequent health and reduced subsequent health services utilization.

The purpose of this study was not only to evaluate whether there was a general effect but also to find and explore any specific effects that payment for CS might have had. Thus, the study emphasizes aspects of utilization, defined by diagnosis, and populations expected to be especially sensitive to such payment.

#### Selection of "At Risk" Populations

The base population for this study was individuals age 65 and over, eligible for both Medicaid and Medicare, not enrolled in managed care and not resident in nursing homes, who filled prescriptions at demonstration pharmacies. The elderly were chosen primarily because the they are, on average, in relatively delicate health and receiving relatively many prescription drugs. Thus, they are at greater risk for drug-related health problems and were hypothesized to be more likely to benefit from cognitive services than younger populations. Dual eligibles were chosen because the intervention was directed at Medicaid recipients, the vast majority of elderly Medicaid recipients are also enrolled in Medicare, and the data would not be consistent if dual eligibles were mixed with those who were only eligible for Medicaid. In addition, four other subpopulations were defined for the analysis:

---

[14] Pharmacies receiving per-service payment did *report* more cognitive services that pharmacies not receiving such payment. Some of this difference may stem from differences in reporting, the remainder from differences in behavior. See Abt Associates, op. cit.

- Dually eligible individuals under age 65;
- Dually eligible elderly with diabetes;
- Dually eligible elderly with three or more prescribers:
- Dually eligible elderly with ten or more prescriptions

Finally, for the analyses of utilization in diagnoses known to be related to certain prescribing problems, the diagnoses were aggregated to the level of the drug class (for example, all diagnoses known to be related to less-than-optimal prescribing of NSAIDs) and the study population was subset to individuals who used drugs in that class. Drug classes were those included in the PSU screener criteria.

### Data Base and File Structure

The data used for this study were the Medicaid drug claims and the Medicaid and Medicare health services utilization claims for individuals in the state of Washington in 1993 (to provide a baseline) and 1994-95. (The demonstration took place between February of 1994 and September of 1995.)

Individualized baseline and follow-up periods were constructed for each sample member. Follow-up periods began on the day that the person first went to an active demonstration pharmacy and lasted until September 1995, loss of eligibility for either Medicaid or Medicare, entry into a nursing home, or death, whichever came first. Similarly, baseline periods ended the day before the person first went to an active demonstration pharmacy and stretched back 12 months or to the first month the person was eligible for both programs or not in a nursing home, whichever came last. These windows were intended to capture the period when the demonstration could and could not have affected health services utilization and when Medicaid and Medicare claims would offer a complete record of health services utilization. The study population included only individuals with positive follow-up time, this being the population that would exhibit intervention effects, and at least three months of baseline because many important covariates were measured in the baseline period. (Results from analyses of those individuals lacking baseline data were broadly similar to those reported here.) The final samples also excluded the relatively few individuals who filled prescriptions in both treatment and control pharmacies.

The elderly population was further subset to groups especially likely to benefit from the demonstration. Diabetics were identified via claims for insulin or oral hypoglycemics at any point in 1993-1995; individuals with multiple prescribers via claims from at least three physicians in a single baseline month; individuals with complex drug regimens via claims for at least ten distinct drugs, as indicated by distinct NDC-9 codes, in a single baseline month.

For some analyses, the population was subset to users of a particular drug, based on use during the baseline period. This guaranteed that the defined subpopulations were independent of intervention effects, but excluded individuals who began using screener drug during the demonstration. This could be a serious omission if, as seems likely, new users are particularly good candidates for adverse drug reactions. Therefore analyses were repeated using an alternative sampling frame that induced individuals who used drugs in the class during either the baseline or the follow-up period. Results for this second population were similar to those for the first and were omitted for brevity.

As Exhibit 4.4.1 shows, beginning with all Medicaid eligible persons who filled prescriptions at any demonstration pharmacy between 1993 and 1997 (254,032), analysts eliminated those with inadequate baseline or follow-up data (106,706), those who could not have been exposed to the demonstration (51,812) and those who became 65 during the study (1,687). Individuals who filled prescriptions in both treatment and control pharmacies (4,120) were also eliminated, leaving a "net usable study population" of 89,707. Of this group, 18 percent (15,889) were dually eligible, and from among these individuals, five analysis samples were created: a) all elderly (7,809), b) elderly diabetics (1,680), c) elderly with three or more prescribers (1,749), d) elderly with ten or more prescriptions (952) and e) non-elderly (8,080).

**Exhibit 4.4.1**

**HCFA DUR Demonstration Evaluation**
**Derivation of Washington Project CARE Downstream Analysis Samples**

| | | |
|---|---|---|
| 1. | Medicaid eligible, filled prescription at demonstration pharmacy, 1993-1997 | 254,032 |
| | *Less* | |
| | Individuals with insufficient data (no follow-up period, baseline less than three months) | (106,706) |
| | Not exposed (timing of first prescription) | (51,812) |
| | Became 65 during study | (1,687) |
| | Filled prescriptions in both arms | (4,120) |
| 2. | Net usable study population | 89,707 |
| | Dually eligible (Medicaid/Medicare) | 15,889 |
| | **Elderly (65+ at beginning of study, 9/30/95)** | **7,809** |
| | **Elderly diabetic** | **1,680** |
| | **Elderly with 3+ prescribers** | **1,749** |
| | **Elderly with 10+ prescriptions** | **952** |
| | **Non-elderly (<65 at end of study, 9/30/95)** | **8,080** |

Source: Washington Project C.A.R.E. Medicare/Medicaid Files

### Utilization Measures

Downstream utilization measures included indicators both of whether or not a service was utilized and, for users, how much, usually in terms of total combined Medicare and Medicaid payments. Exhibit 4.4.2 describes these measures. They include inpatient and professional services, as well as outpatient emergency room admissions. ER measures were expected to capture acute prescription-drug related episodes treated in an outpatient rather than an inpatient setting. [15] The study employed both broadly-defined measures (for example, the frequency of all inpatient admissions in the demonstration sample) and more narrowly-defined measures ( the frequency of inpatient admissions for diagnoses linked to less than optimal prescribing of specific drugs). To support these more targeted measures, researchers at Penn State University assembled a data base of known relationships between clinical outcomes and selected prescribing problems, namely those problems that are included in the University of Maryland/Philadelphia College of Pharmacy Science screener. This screener covers four prescribing problems — dosage, drug/drug interactions, duration of therapy, and therapeutic duplication — and eight classes of drugs — ACE inhibitors, anti-depressants, anti-psychotics, benzodiazepines, digoxin, calcium channel blockers, H2RAs, and NSAIDS.

**Exhibit 4.4.2**

**HCFA DUR Demonstration Evaluation**
**Utilization Measures: Washington Project CARE Downstream Analysis**

| Measure | Description |
|---|---|
| Inpatient (0,1) | Probability of an inpatient admission reimbursed by Medicare or Medicaid |
| Inpatient ($) | Total Medicare and Medicaid payment, including professional services, for hospital admissions |
| Professional (0,1) | Probability of a professional service encounter (single day's interaction with a professional provider) delivered in a non-hospital setting |
| Professional ($) | Total Medicare and Medicaid payment for professional service encounters |
| ER (0,1) | Probability of an outpatient emergency room encounter |
| ER($) | Total Medicare and Medicaid payment for outpatient emergency room encounters |
| Total (0,1) | Probability of any Medicare or Medicaid service* |
| Total ($) | Total Medicare and Medicaid payment for all services* |

*Includes inpatient, professional, emergency room, and skilled nursing facility services
Source: Washington Project C.A.R.E. Medicare/Medicaid Files

---

[15] Analysts studied skilled nursing facility utilization, and measures of total utilization and expenditure include SNF services. However, SNF admissions in all of the analysis populations were infrequent (1.6 percent among all dually eligible elderly sample members), and absent altogether for some of the smaller samples. SNF utilization was, therefore, not included as a separate dependent variable.

The PSU outcomes data base is comprehensive and well-documented. It is the result of a search of primary literature via electronic bibliographic data bases, a manual search of selected secondary sources such as the major drug compendia and resource texts, and extensive review and comment from clinical experts. However, it is important to remember that this data base only covers *selected* prescribing problems, while the Washington demonstration potentially touched on *all prescribing* problems.[16] Furthermore, the data base (and OPDUR more generally) are only concerned with drugs and prescriptions while cognitive services may influence not only drugs and prescriptions but also patients' knowledge and behavior regarding prescriptions. Within that data base, targeting diagnoses were selected for this study if both the diagnosis and the associated prescribing problem were sufficiently prevalent in the Washington Medicaid population that the study would have sufficient power to detect a demonstration effect of moderate size (in the neighborhood of one quarter of one standard deviation).

### Independent Variables

In modeling the effects of the demonstration on downstream utilization, analysts first developed measures to characterize demonstration exposure and then selected, from the available data, other covariates that theory suggests may have utilization impacts separate from those due to exposure. The independent variables selected for this model are listed in Exhibit 4.4.3.

### Characterization of Demonstration Exposure

Although this demonstration was designed as a random experiment, randomization occurred at the level of the pharmacy while many of the benefits of payment for cognitive services were expected to accrue at the level of the pharmacy patron/Medicaid recipient. This posed several modeling challenges. First, individuals had the potential to go to multiple pharmacies and, potentially, to go to a mix of intervention, control, and non-demonstration pharmacies. The study adopted a simple approach. Individuals were considered members of the intervention group if they ever went to an intervention pharmacy and never went to a control pharmacy. They were considered members of the control group if they ever went to a control pharmacy and never went to an intervention pharmacy. Individuals who went to both intervention and control pharmacies were excluded from the study.

In order to calculate and apply the cluster-level averages, it was necessary to assign individuals to clusters. This assignment was based on the cluster affiliation of the first active demonstration pharmacy to which the person went.

The intervention/control indicator was admittedly a simplification of a potentially complex exposure. Individuals deemed members of the intervention group varied in terms of the number of prescriptions they filled at intervention pharmacies, the complexity of their drug regimens, the length of time they

---

[16] One type of less-than-optimal prescribing is the presence of problem prescriptions; the discussion in this paper emphasizes this type of problem prescribing because this is where cognitive services are most likely to have an impact. However, the absence of potentially beneficial prescriptions is another type of less-than-optimal prescribing and is probably of equal importance.

**Exhibit 4.4.3**

**HCFA DUR Demonstration Evaluation**
**Independent Variables: Washington Project CARE Downstream Analysis Samples**

| Variable | Definition |
|---|---|
| Demonstration exposure | |
| Treatment (0,1) | Demonstration status of individual: treatment if all prescriptions during study period filled at treatment pharmacies, control if all filled at control pharmacies |
| Other independent variables | |
| Age | Nine qualitative variables: 20-29, 30-39, 40-49, 50-59, 60-64, 65-69, 70-79, 80-84, 85+ |
| Sex | |
| Race | Three qualitative variables: white, black, other |
| Eligibility | Medicaid eligibility - blind or disabled |
| Baseline utilization | Value of the dependent variable during baseline period |
| DCG/HCC | Predicted annual Medicare payments |
| Baseline drugs | Maximum number of NDC-9 level distinct drugs in a single month during baseline period |
| Length of baseline period (months) | |
| Length of follow-up period (months) | |
| Wave | Recruitment wave of pharmacy filling individual's first prescription (three waves in Washington) |

Source: Washington Project C.A.R.E. Medicare/Medicaid Files

went to intervention pharmacies, and whether their exposure was relatively late in the demonstration (when the gap in documented rates of CS between intervention and control pharmacies was at its widest.) One might expect the intervention's impact to be increasing in all of these dimensions. Members of the intervention group also differed in age, and a hypothesis of the evaluation was that demonstration effects would be especially pronounced among the elderly. To capture this complexity, the study experimented with alternative specifications featuring interactions of intervention status with appropriate variables (always also incorporating both the intervention status indicator and the variable itself). No consistent significant results emerged, and results of these interactions aren't shown.

A second modeling issue was that the regression framework required adjustment in order to account for this experimental structure. The number of pharmacy clusters was fairly small (37) leading to differences in measured demographics and baseline outcome rates between intervention and control recipients. While one could control for any measured differences directly within the regression framework, the presence of these differences raised the suspicion that there could also be unmeasured differences in demographics, access, or medical practice across clusters that would be correlated with intervention status and also with some or all of the dependent variables. Because the number of clusters was small relative to the number of recipients, these differences might not average out between intervention and control groups despite the large number of recipients leading to misleading estimates of intervention effects unless an appropriate adjustment was made. Specifically, traditionally-estimated standard errors would be smaller than the true standard errors.

The first adjustment was to include the cluster-level averages and demographics discussed above. These variables controlled directly for measurable differences among individuals and actually improved the precision of estimated intervention effects. The second was to employ the econometric technique known as variable effects; this technique effectively adjusted the standard errors of estimated parameters to account for the number of clusters and the amount of the marginal variation in the dependent variable that they explained.

### Other Covariates

The Andersen/Newman model of factors explaining health service utilization suggests three sets of independent variables [Anderson & Newman (1973)]. Predisposing factors include socio-demographic factors, here sex, race, and age. Each model also included an indicator of Medicaid blind/disabled eligibility. Enabling factors include measures of access to health care, both financial and geographic. All members of the sample had roughly equal financial access to care, given that they were dually eligible for Medicare and Medicaid. Direct measures of geographic access were not available, but each regression included the average value of the dependent variable during the baseline period for the individual's pharmacy cluster, effectively the individual's medical community defined via pharmacies with common prescribers. This variable served as a proxy for the net effect of geographic factors enhancing or reducing access to care and for geographic variation in the practice of care, such as variations in coding of conditions and diagnoses and in treatment protocols.

The third set of factors in the Andersen/Newman model are measures of need, such as diagnosis. For this study, the objective was to control for need, exclusive of the need that was influenced by the demonstration. The principal measure of need was based on the Diagnostic Cost Group (DCG) and Hierarchical Coexisting conditions (HCC) models and software developed by Ellis, Pope, Iezzoni, and others.[17] In short, the model used in this paper (the HCCPH-prospective model) reads information on a person's age, sex, diagnoses, life sustaining medical procedures, and selected hospitalizations during the baseline period and uses that information to predict Medicare payments for that person in the next year. As an additional measure of need, regressions controlled for the maximum number of distinct drugs, (as measured by NDC-9s), that the individual received during a single month of the baseline period.

The lengths of each person's follow-up and baseline periods were two other controls. Clearly, the longer the follow-up period, the greater that the recipient's expected health services utilization. The longer the baseline period, the more complete the record of the individual's prior health services utilization and, by proxy, ongoing medical problems. One would expect that the measure of expected health spending would be biased downward for some individuals with short baseline periods and therefore that, other things equal, a short baseline would be correlated with higher utilization in the follow-up period.

---

[17] These models are described more fully in R. P. Ellis. G. C. Pope, L. Iezzoni et al, "Diagnosis-Based Risk Adjustment for Medicare Capitation Payments, *Health Care Financing Review*, Spring, 1996.

Finally, as participating pharmacies were recruited in waves rather than at a single point in time, each person was assigned an indicator that identified the recruitment wave of the pharmacy that filled that person's initial prescription. Hypothesizing a wave effect implies a belief in the possibility of a "halo" effect, in which a pharmacy's demonstration participation itself, independent of access to CS payment, may affect CS behavior. Two competing hypotheses about how wave affects utilization might be entertained. On the one hand, late-wave treatment and control pharmacies might be better informed than early-wave pharmacies about the demonstration, leading to increased CS activity and, presumably, reductions in downstream utilization in both arms. On the other hand, the time to observe utilization is foreshortened for late-wave recruits, suggesting that, even after adjustment for length of follow-up, utilization effects will be less frequent for late than for early wave recruits.

### Regression Specification

The model used to estimate demonstration effects was specified in the following way:

$$y_{ijt_f} = \alpha + \beta_1 tr_{ij} + X_i \beta_2 + Z_i \beta_3 + \beta_4 y_{jt_b} + \mu_j + \nu_{ij}$$

with:

| | |
|---|---|
| $y_{ijtf}$ | utilization of recipient i associated with cluster j during the follow-up period |
| $tr_{ij}$ | indicator that recipient i filled a prescription at treatment pharmacies and not at control pharmacies during the demonstration period |
| $X_i$ | predisposing (socio-demographic) and need characteristics of recipient i |
| $Z_i$ | length of follow-up and baseline windows for recipient i |
| $y_{jtb}$ | cluster average of outcome variable in baseline period |
| $\mu_j + \upsilon_{ij}$ | error |

The compound error term is the structure required for the variable effects model. As noted above, models using both the usual single-component error and compound error terms were tested, showing no differences in sign or significance of estimates of treatment effects. Findings below are based on the single-component error model.

Models predicting downstream utilization were estimated for:

- Component and total utilization measures, for each of the five populations/subpopulations;

- Selected probabilities (of inpatient admissions, of professional services, of all services) and total payments for all services, for services with diagnoses related to the misuse of each of seven drug categories (ACE inhibitors, antidepressants, antipsychotics, benzodiazepines, calcium channel blockers, digoxin, NSAIDs, total targeted)[18], for each of the five populations/subpopulations

### 4.4.3 Findings

#### Descriptive Statistics

Covariate means are useful 1) for comparing the principal analysis population (dually eligible elders) with selected subpopulations and 2) for showing that, though pharmacy assignment in this demonstration was random (in clusters), randomization did not succeed in minimizing differences between treatment and control populations. Exhibit 4.4.4 shows the means of the variables included in the regressions, for all five analysis populations used in the study.[19] Subpopulations were slightly

**Exhibit 4.4.4**

**HCFA DUR Demonstration Evaluation**
**Washington Project CARE Downstream Analysis Sample Characteristics**

| Measure | | | Population | | |
|---|---|---|---|---|---|
| | All elderly (n=7,809) | Non-elderly (n=8,080) | Diabetics (n=1,680) | 3+Prescribers (n=1,749) | 10+Prescriptions (n=952) |
| Age(30-39) | -- | 0.32 | -- | -- | -- |
| Age(40-49) | -- | 0.28 | -- | -- | -- |
| Age(50-59) | -- | 0.20 | -- | -- | -- |
| Age(60-64) | -- | 0.04 | -- | -- | -- |
| Age(70-79) | 0.44 | -- | 0.45 | 0.44 | 0.46 |
| Age(80-84) | 0.14 | -- | 0.12 | 0.12 | 0.11 |
| Age(85+) | 0.12 | -- | 0.07 | 0.07 | 0.06 |
| Race(black) | 0.05 | 0.06 | 0.09 | 0.06 | 0.05 |
| Race(other) | 0.24 | 0.08 | 0.26 | 0.17 | 0.16 |
| Sex(male) | 0.26 | 0.50 | 0.25 | 0.23 | 0.21 |
| Blind/disabled | 0.13 | 0.96 | 0.15 | 0.15 | 0.15 |
| DCG/HCC | $4,629 | $3,948 | $6,988 | $6,332 | $8,372 |
| Baseline drugs | 2.6 | 2.5 | 5.0 | 7.7 | 12.2 |

Source: Washington Project C.A.R.E. Medicare/Medicaid Files

younger than the dually-eligible elderly population from which they were drawn; from six to seven percent were 85 or over, compared to 12 percent in the elderly population. African Americans and other non-white minorities were more prevalent among elderly diabetics. Other minorities were less prevalent among non-elderly and elderly with multiple prescribers or prescriptions. Gender differences between elderly (26 percent male) and non-elderly (50 percent male) are large in the expected direction. Elderly men are less frequently represented in the multiple prescription and prescriber samples. Expected eligibility differences between elderly and non-elderly reflect programmatic rules, not changes in acuity or chronicity of conditions with aging. Finally, predicted costs and use of multiple baseline drug measures clearly distinguish the greater health services needs of subpopulations. DCG/HCC predicts annual costs for the subpopulations from 50 to 80 percent above average costs for the whole elderly population. Subpopulation members use from two to four and a half times the number of drugs used by the average elderly population member in a month.

---

[18] Statistical power was insufficient to conduct tests of the effects of CS payment on downstream utilization associated with H2 RA drugs.

[19] Certain qualitative variables excluded from the multivariate analyses are not shown: age (20-29) for the non-elderly, age (65-69) for the elderly, race (white) for all samples.

As has been stated, the design of the demonstration involved random assignment of geographic clusters of pharmacies to the intervention and control arms. Exhibit 4.4.5 shows the means of selected variables for intervention and control recipients separately. Despite random assignment of clusters, there were some differences between the two groups. Treatment members were slightly younger than control group members. Higher percentages of individuals in the treatment group were non-white (black and other racial minorities). Furthermore, levels of drug utilization, mean inpatient hospitalization payments, and mean payments for professional services were higher among intervention recipients suggesting that they were in poorer health.[20] Possibly, these differences resulted from the fact that there were demographic differences among geographic areas in Washington, that the number of pharmacy clusters was rather small, and, thus, random assignment of clusters did not result in entirely comparable intervention and control groups despite the large sample size. The regression specifications controlled for these measured characteristics, and the variable effects model adjusts the standard errors to take account for any residual explanatory power associated with the cluster.

**Exhibit 4.4.5**

**HCFA DUR Demonstration Evaluation**
**Washington Project CARE Downstream: Dually Eligible Elderly Treatment and Control Recipient Characteristics**

| Measure | Treatment (N = 4,591) | Control (N = 3,218) |
|---|---|---|
| Age*** | 74.5 years | 75.0 years |
| Race(black)*** | 0.07 | 0.03 |
| Race(other)** | 0.25 | 0.23 |
| Sex(male) | 0.26 | 0.27 |
| DCG/HCC | $6,117 | $6,068 |
| Baseline drugs** | 5.2 | 5.0 |
| Baseline months | 11.0 | 10.8 |
| Follow-up months | 13.1 | 12.5 |

* The difference between the arms is statistically significant at the 10 percent level
** The difference between the arms is statistically significant at the 5 percent level
*** The difference between the arms is statistically significant at the 1 percent level
Source: Washington Project C.A.R.E. Medicare/Medicaid Files

---

[20] It might also suggest that the prevailing style of medical practice was more resource-intensive in intervention areas.

### Multivariate Analyses

#### *Model Performance*

Though the analyses provided no consistent support for the hypothesis that CS payment reduces downstream utilization, models used in the analyses generally performed well, given the limited availability of data for measuring covariates. Most models explained from 5 to 10 percent of variation in the dependent variables. Signs and significance levels of some model covariates were consistent across many of the models. For example, measures of baseline medical need, including predicted annual costs, maximum number of drugs in a month, and baseline measures of the dependent variables, were inevitably positive and highly significant. The number of follow-up months was also positive and significant, as expected. In addition, individuals who filled their first prescriptions in pharmacies recruited later into the study generally had lower probabilities and smaller levels of downstream utilization, independent of the lengths of their follow-up periods. Demographic variables were less likely to generate consistently significant estimates, but the estimated directions of effects (signs) were reasonably consistent for some measures. For example, African Americans were less likely to experience an inpatient admission than others. Age measures generally depicted increasing frequency of utilization with advanced age, but decreasing intensity. For example, in a model that predicts inpatient admissions for dually-eligible elderly, all three age-defined coefficients (70-79, 80,-84, and 85+) were positive, they increased in absolute size from the youngest to the oldest, and two were highly significant. Coefficients in a model predicting the costs of those individuals admitted for an inpatient stay were uniformly negative, and suggestive of some tendency for a reduction in costs for the oldest category.

#### *Predictions of Total Utilization: All Study Populations*

CS payment has no consistent effects on components and aggregate measures of downstream utilization.

Exhibit 4.4.6 shows the signs and levels of significance for treatment variables in models that predict inpatient, professional services, emergency room and total utilization in the five study populations/ subpopulations. The exact numerical values of coefficients and their associated p-values provide no useful information, in the absence of any meaningful patterns in the estimates.

Few rows or columns in this exhibit show consistency of signs, and there are no instances in which the direction of effects, for any population or measure of utilization, is always negative. Only the non-elderly (adult disabled) show more than one statistically significant treatment effect; two of the three (professional services payments and the probability of an ER admission) are negative, while ER payments are positive. In fact, ER payments are positive across all populations, and significant for dually eligible elderly and non-elderly individuals. Given the ambiguity of the other estimates in this table, this finding suggests less an effect of the demonstration than a defect in the design, perhaps the influence of unmeasured characteristics of treatment individuals that make them more costly than controls to treat in an ER setting.

**Exhibit 4.4.6**

**HCFA DUR Demonstration Evaluation**
**Regression Adjusted Estimates of Total Utilization Effects: All Analysis Populations**

| Utilization Measures | Populations | | | | |
|---|---|---|---|---|---|
| | All elderly | Non-elderly | Diabetics | 3+Prescribers | 10+ Prescription |
| Inpatient (0,1) | (+) | (-) | (-) | (+) | (-) |
| Inpatient ($) | (-) | (+) | (-) | (-) | (-) |
| Professional services (0,1) | (+) | (+) | (+) | (-) | (+) |
| Professional services ($) | (-) | (-)* | (-) | (+) | (+) |
| ER (0,1) | (-) | (-)* | (+) | (+) | (-) |
| ER ($) | (+)*** | (+)* | (+) | (+) | (+) |
| Total (0,1) | (+) | (+) | (+) | (-) | (+) |
| Total ($) | (+) | (-) | (-) | (+) | (+) |

* The difference between the arms is statistically significant at the 10 percent level
** The difference between the arms is statistically significant at the 5 percent level
*** The difference between the arms is statistically significant at the 1 percent level
Source: Washington Project C.A.R.E. Medicare/Medicaid Files

***Predictions of targeted utilization*: *all elderly, elderly diabetics, elderly with 10 or more prescriptions*.** There is mixed evidence for treatment "effects" on targeted utilization. Exhibit 4.4.7 shows signs and significance levels for inpatient and professional services utilization associated with seven drug categories, separately and aggregated, for all dually eligible elderly, elderly diabetics and elderly with ten or more prescriptions. Hypothesized negative signs on inpatient and professional utilization appear consistently across the three populations for calcium channel blockers; statistically significant estimates are found for elderly diabetics and multiple drug users. Most of the treatment coefficients for digoxin are negative, of which two are significant at the ten percent level, but, counter to intuition, the most statistically significant coefficient ( measuring the probability of any professional services among all elderly) is positive. In fact, three of the five coefficients that are significant at a five percent level are positive.

### 4.4.4 Discussion

Despite the appearance of "patterns" of significant negative estimates in the analyses of targeted utilization, it would be imprudent to conclude that these findings demonstrate CS payment effects. The PSU Screener drugs were never incorporated into Project C.A.R.E.'s interventions. Washington pharmacists who participated in Project C.A.R.E. were given no special tools or training to focus their attention on the PSU drug categories. The evaluator analyzed these drug categories in both

**Exhibit 4.4.7**

## HCFA DUR Demonstration Evaluation
## Regression Adjusted Estimates of Drug/Diagnosis Specific Utilization Effects: All Elderly, Elderly Diabetics, Elderly With 10+ Prescriptions

| Drug Categories/ Utilization Measures | Populations | | |
|---|---|---|---|
| | All Elderly (N=5,324) | Elderly Diabetics (N=1,345) | Elderly/10+ Prescriptions (N=917) |
| *Ace Inhibitors* | | | |
| Inpatient (0,1) | (+)** | (-) | (+) |
| Professional services (0,1) | (-) | (-) | (-) |
| *Antidepressants* | | | |
| Inpatient (0,1) | (+) | (-) | (+) |
| Professional services (0,1) | (+) | (+) | (+) |
| *Antipsychotics* | | | |
| Inpatient (0,1) | (-) | (-) | (-) |
| Professional services (0,1) | (-) | (-)* | (-) |
| *Benzodiazepines* | | | |
| Inpatient (0,1) | (+) | (+) | (-) |
| Professional services (0,1) | (+) | (-) | (+) |
| *Calcium Channel Blockers* | | | |
| Inpatient (0,1) | (-) | (-)** | (-)** |
| Professional services (0,1) | (-) | (-)* | (-) |
| *Digoxin* | | | |
| Inpatient (0,1) | (-)* | (-) | (-) |
| Professional services (0,1) | (+)** | (-) | (-)* |
| *NSAIDs* | | | |
| Inpatient (0,1) | (-) | (-) | (+) |
| Professional services (0,1) | (+) | (+) | (+)** |
| *All targeted (0,1)* | | | |
| Inpatient (0,1) | (-) | (-)** | (-)* |
| Professional services (0,1) | (+) | (-) | (-) |
| *All targeted ($)* | | | |
| Inpatient ($) | (-) | (-) | (+) |
| Professional services ($) | (+) | (-) | (-) |

* The difference between the arms is statistically significant at the 10 percent level

** The difference between the arms is statistically significant at the 5 percent level

*** The difference between the arms is statistically significant at the 1 percent level

Source: Washington Project C.A.R.E. Medicare/Medicaid Files

Washington and Iowa for two reasons. First, because these drug categories currently provide the only well-documented linkages between drug misuse and utilization markers defined by diagnosis, it seemed reasonable to explore potential demonstration effects at this more disaggregated level. This approach is refined even further in sections of this chapter that assess effects on specific clinical outcomes (Sections 4.7 through 4.11). Unfortunately, despite the appearance of significant negative treatment coefficients for the overall use of professional services or for utilization associated with calcium channel blockers,[21] the results of these analyses for Washington cannot stand alone as evidence for CS effects. Second, these results provide a context for assessing similar analyses in Iowa, where the OPDUR system incorporated criteria for assisting Iowa pharmacists in evaluating these specific drug groups. In comparison to these results for Project C.A.R.E., one should expect to find more, and more significant, negative treatment effects in the PSU drug categories in Iowa. These comparisons are discussed below in Section 5.6.

## 4.5 Effects of Project C.A.R.E on Prescription Use and Costs

This section presents findings of the effect of Project C.A.R.E. on prescription drug use and costs. The objective of this aspect of the DUR Demonstration evaluation was to determine whether there were statistically significant reductions in drug cost and use between Medicaid recipients who filled their prescriptions in demonstration pharmacies relative to pharmacies in the control group during the demonstration period. The premise of the study is that patients who receive optimal drug therapy will generally use fewer medications overall and that drug costs will also be lower as a result. Given this premise, it was hypothesized that to the extent that cognitive services improve patient drug therapy, then both Medicaid drug use and costs should be lower for patients who visit demonstration pharmacies compared to the controls. Since the demonstration was not targeted to specific drug groups (unlike the Iowa OPDUR Demonstration Project) the hypothesis was tested for all drugs combined and for various therapeutic categories including the nine most costly drug groups (based on Washington Medicaid outlays in 1995) and the eight classes of drugs screened by the PSU screener.

The study design closely follows that presented in the previous section on downstream medical utilization and costs. For this reason, the discussion of data and methods is quite brief consisting primarily of a summary of study procedures. As in the case of the downstream analyses, no evidence of a demonstration effect was found in any of the five subpopulations reviewed. Regression results are presented for the dual-eligible 65+ age group. Results from the other four population groups are briefly discussed.

### 4.5.1 Data and Study Population

This aspect of the study utilized the same analytic dataset as for the downstream analyses reported in Section 4.4. The study population was also the same, consisting of dual-eligible (Medicare/ Medicaid)

---

[21] Patterns of documented CS activity by drug category do not bear any obvious relationship to these findings. Although Project C.A.R.E. treatment pharmacies provided more CS per 100 prescriptions for calcium channel blockers than control pharmacies (1.17, compared to 0.36) similar large differences were observed for drugs with no significant treatment effects. For other drugs (e.g., digoxin and antipsychotics) the average CS rate in treatment pharmacies was even higher (2.99 and 2.15), and treatment and control differences were similarly large and highly significant. (Reported in Christensen et al., 1996: pages 49, 50)

Medicaid recipients who met the following inclusion and exclusion criteria: (1) eligible for both Medicare and Medicaid during the baseline and follow-up periods, (2) exposed to the demonstration, and (3) did not fill prescriptions in both a demonstration and control pharmacy (see Exhibit 4.4.1). Application of these criteria resulted in a sample of 7,809 elderly subjects and 8,080 non-elderly subjects. The elderly sample was further subset into three at-risk groups: elderly diabetics (1,680), elderly with three or more prescribers (1,749) and elderly with 10 or more prescriptions (952). The sample selection process is described more fully in Section 4.4.2.

### 4.5.2 Methods

The statistical methods employed to test the study hypotheses are identical to those used in the downstream analysis (see Section 4.4.2). Multiple regression models were specified to produce conditional estimates of the demonstration impact on a wide range of drug use and cost outcomes holding constant other factors that might contaminate or bias the demonstration coefficients. The dependent variables in these models included total prescriptions, total Medicaid drug payments, individual totals for the most commonly used drug categories (central nervous system agents, gastrointestinal drugs, cardiovascular drugs, analgesics and anesthetics, endocrine and metabolic drugs, respiratory agents, anti-infectives, neuromuscular drugs, and topical products), and individual totals for each of the eight PSU drug classes (ACE inhibitors, calcium channel blockers, digoxin, antidepressants, antipsychotics, benzodiazepines, H2RAs, and NSAIDs). In all, 270 multiple regression models were estimated.

Except for the choice of dependent variables, model specifications were identical to those used in the downstream analyses. The independent variables are listed in Exhibit 4.4.3. Demonstration effects on the probability of any prescription drug use in the follow-up period were estimated with Probit. Single-component error OLS models were used to estimate demonstration effects on counts and costs of drugs used by those subjects with at least one prescription fill in the follow-up period. Because the distributions of drug cost and use are right-skewed, the dependent variables in these regressions were transformed into log form. The coefficients in these models can thus be read as elasticities.

Test regressions with interaction terms involving the demonstration effect variable failed to produce significant results. The same was true for regressions estimated using a variable effects specification. These results are not reported.

### 4.5.3 Findings

The demographic characteristics of the five study populations are summarized in Exhibits 4.4.4 and 4.4.5 in the downstream analysis section. As can be seen there, the study subjects are predominantly white females in both entitlement categories (aged and blind/disabled), but there are other major differences between the two groups. Individuals entitled through disability or blindness are much younger (modal age, 30-39), much more likely to be white (84% versus 71% for the aged), and have lower predicted resource use according to the DCG/HCC index ($3,948 versus $4,629). The "at risk" elderly subpopulations are somewhat younger but much sicker than the elderly sample as a whole (the DCG/HCC index for elderly with 10+ prescriptions is nearly double that for the whole). As noted in the previous section, the control and treatment groups differ significantly in selected domains (age, race, and baseline drug use) thereby necessitating the statistical control offered by multiple regression techniques.

Exhibit 4.5.1 presents means and standard deviations for the dependent variables on behalf of the 65+ population sample. The top panel of the table reports utilization and cost statistics for the nine high-expense drug groups. The bottom panel reports utilization and cost in the eight PSU screener drug categories. Totals for all drug use are shown in the bottom row. By definition, all subjects filled at least one prescription in the follow-up period (evidence of drug use was necessary for assigning subjects to the treatment or control group). The average elderly recipient utilized 24 prescriptions or refills during the 13 months of the follow-up period at an average cost of $633. All of the elderly at-risk subpopulations had much higher usage rates. Elderly diabetics filled an average of 43 scripts in the follow-up period. Recipients with three or more prescribers filled 50 prescriptions, while those with 10 or more prescriptions filled an average of 70. Average drug costs were also much higher for these subpopulations, ranging from $1,236 for elderly diabetics to $2,100 for the 10+ prescription group. Dual-eligible recipients under age 65 averaged 25 prescriptions at a cost of $721 during the follow-up period.

The distribution of drug use by therapeutic class shown in Exhibit 4.5.1 followed an expected pattern. The elderly used more cardiovascular drugs than any other type of prescription medication (71 percent of the sample used these drugs at a mean cost of $219 in the follow-up period). More than 60 percent of the sample filled one or more prescriptions for analgesics/anesthetics and anti-infectives, and except for neuromuscular drugs, more than 40 percent filled prescriptions in the other high-expense drug classes. Among the PSU screener drugs, utilization rates were highest for NSAIDs, calcium channel blockers, and H2Ras. Relatively few elderly recipients took antipsychotics (ten percent) or benzodiazepines (six percent).

Exhibit 4.5.2 presents multivariate results of the hypothesis test that the C.A.R.E project reduced drug use and cost among the elderly dual-eligible sample. The cells in this table report coefficient values and standard errors for the treatment/control dummy variable from the 30 regressions relating to total drug use and drug use within the high-expense therapeutic categories. It is clear from this table that there was no treatment effect. Only four of the coefficients reached conventional levels of statistical significance and three of these were in the opposite direction (+) hypothesized. Indeed only seven of the 30 coefficients were in the hypothesized direction (-). Results from the PSU screener drug regressions also showed no treatment effect. The same was true for the regression models estimated for the at-risk elderly subpopulations and for the under 65 dual-eligible sample. In short, the C.A.R.E project made no difference in how many recipients used drugs, how many they used, or how much it cost the state to pay for them.

As in the case of the downstream analyses, the regression models themselves performed relatively well in terms of overall explained variance. The coefficients of determination ($R^2$) in the OLS regressions (adjusted for number of covariates) ranged from seven to 35 percent in the therapeutic class regressions and between 58 and 77 percent in the total drug use regressions. As expected, the two variables with the greatest predictive power were the number of follow-up months (the longer the observation period the greater the use and cost) and the maximum number of GPI classes represented in the baseline period (also positive). Other explanatory variables that generally reached statistical significance were the HCC/DCG cost predictor (+), sex (women use more drugs), other race (persons other than white or black use fewer drugs), and age (-). These findings are similar to those reported in the downstream analyses (Section 4.4.3) and the explanations offered there apply here as well.

**Exhibit 4.5.1**

## HCFA DUR Demonstration Evaluation
## Characteristics of Study Sample Recipients' Drug Utilization by Therapeutic Class: Dual-eligible 65+ (n=7,809)

| | Sample Statistics | | | | | |
|---|---|---|---|---|---|---|
| | Percent Using Drugs | | Count of Drugs Used | | Cost of Drugs Used | |
| Drug category (high expense) | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| Central nervous system agents | 42.6% | 0.49 | 4.5 | 3.2 | $66.02 | $4.66 |
| Gastrointestinal agents | 52.0 | 0.50 | 4.4 | 3.0 | 132.95 | 5.99 |
| Cardiovascular drugs | 71.4 | 0.45 | 9.6 | 3.2 | 219.20 | 4.71 |
| Analgesics and anesthetics | 63.9 | 0.48 | 4.2 | 3.0 | 59.70 | 4.81 |
| Endocrine and metabolic drugs | 44.4 | 0.50 | 5.4 | 2.9 | 93.69 | 4.57 |
| Respiratory agents | 43.2 | 0.50 | 3.9 | 3.4 | 61.56 | 5.47 |
| Anti-infectives | 62.6 | 0.48 | 2.6 | 2.3 | 46.99 | 3.67 |
| Neuromuscular drugs | 15.6 | 0.36 | 3.5 | 3.2 | 64.72 | 5.10 |
| Topical products | 47.6 | 0.50 | 2.8 | 2.6 | 49.40 | 3.39 |
| Drug category (PSU classes) | | | | | | |
| ACE Inhibitors | 22.9% | 0.42 | 5.3 | 2.6 | $200.34 | $310 |
| Antidepressants | 23.8 | 0.43 | 4.0 | 2.9 | 68.72 | 4.62 |
| Antipsychotics | 9.8 | 0.30 | 2.9 | 3.0 | 51.42 | 3.86 |
| Benzodiazepines | 6.3 | 0.24 | 2.9 | 2.7 | 35.87 | 3.49 |
| Calcium channel blockers | 29.3 | 0.45 | 5.8 | 2.6 | 270.43 | 3.03 |
| Digoxins | 17.8 | 0.38 | 3.5 | 2.3 | 33.12 | 2.41 |
| H2RAs | 29.3 | 0.46 | 3.7 | 2.7 | 275.89 | 3.22 |
| NSAIDs | 32.2 | 0.47 | 2.9 | 2.6 | 83.10 | 4.35 |
| All drugs | 1.00 | 0.00 | 24.3 | 3.6 | $632.70 | $4.62 |

Source: Washington Project C.A.R.E. Medicaid Prescription Drug Files, 1993-96.

**Exhibit 4.5.2**

**HCFA DUR Demonstration Evaluation**
**Regression Results of the Impact of the Washington C.A.R.E. Demonstration on the Probability of any Drug Use, Counts of Drugs, and Cost of Drugs Used by Therapeutic Class: Dual-eligible 65+**

| | Regression Coefficients (standard Errors) for the Demonstration Effect | | |
|---|---|---|---|
| **Drug Category** | **Probability of Use (Probit)** | **Count of Drugs (ln)** | **Cost of Drugs (ln)** |
| Central nervous system agents | 0.01 (0.03) | 0.00 (0.04) | -0.02 (0.05) |
| Gastrointestinal agents | 0.04 (0.03) | 0.03 (0.03) | -0.00 (0.06) |
| Cardiovascular drugs | 0.01 (0.03) | 0.03 (0.03) | 0.02 (0.02) |
| Analgesics and anesthetics | -0.05 (0.03) | 0.04 (0.03) | 0.07 (0.04) |
| Endocrine and metabolic drugs | -0.05 (0.03) | *-0.08 (0.03) | -0.09 (0.05) |
| Respiratory agents | *0.08 (0.03) | 0.03 (0.04) | 0.02 (0.06) |
| Anti-infectives | 0.01 (0.03) | **0.06 (0.02) | 0.04 (0.04) |
| Neuromuscular drugs | 0.02 (0.04) | 0.04 (0.06) | 0.08 (0.09) |
| Topical products | *0.07 (0.03) | -0.00 (0.03) | 0.00 (0.04) |
| All drugs | 0.16 (0.13) | 0.03 (0.02) | 0.03 (0.03) |

Source: Washington Project C.A.R.E. Medicaid Prescription Drug Files, 1993-96.
Notes: Statistical significance indicated by * if p<.05 and ** if p<.01 in two-tailed tests.

### 4.5.4 Discussion

The analysis presented in this section offers no support for the hypothesis that the Washington C.A.R.E. demonstration had a measurable impact on either drug use or cost. The regression results fail to show any pattern of downward trends attributable to the demonstration in any of the five population groups for any of the 18 drug classes studied.

It is possible that different results would have been obtained had the study focused on different population groups (e.g., women and children enrolled in AFDC) or other drug classes. But this seems very unlikely. Dual-eligible recipients tend to be among the sickest members of the Medicaid population and the highest users of prescription drugs. Were payment for cognitive services to have any impact on drug use and cost, one would expect to find it in this group of recipients. Failure to find an impact implies either that pharmacists' cognitive services were ineffectual in controlling pharmaceutical use or that the payment incentives were not structured in a way that engendered utilization-reducing behavior. The C.A.R.E Demonstration Final Report [(Christensen, D. and Holmes, Garth, 1996)] suggests that the latter interpretation may be closer to the mark. While payment for cognitive services appears to increase the number of services provided (or at least reported) the most common types of service rendered (counseling patients, providing case management) would not, in themselves, result in any reduction in use. The type of services that would likely have this effect (unnecessary drug, duplicative therapy, drug-drug interact) comprised fewer than five percent of the cognitive services reported by pharmacists in the treatment group. (Christensen, D. and Holmes, Garth, 1996, p.39). In other words, if the goal is to reduce drug cost and use, then the policy of paying pharmacists to provide additional pharmaceutical care must be targeted to that end.

## 4.6 Effects of the Cognitive Services Demonstration on Screen Failures

The Washington Project C.A.R.E. program was a demonstration that offered payments to pharmacists for providing cognitive services. Payment for cognitive service was intended to provide appropriate financial compensation for the time and effort needed to provide cognitive services. The demonstration was conducted in Washington State from February 1994 to September 1995. The population of interest in this study was Washington residents who were eligible Medicaid, and who filled a prescription at a demonstration pharmacy at least once. This chapter assessed the effects of the Project C.A.R.E. program on the occurrence of screening failures using the drug therapy criteria embedded in the PSU Screener.

While the Washington cognitive services demonstration was not directly targeted toward identifying the same prescribing problems that are the basis for the UM/PCPS criteria, it could nevertheless be argued that an effective cognitive services system should detect and address many, if not all, of the same problems that are described in the UM/PCPS criteria. Further it would be expected that, as a result of the demonstration, physicians would become more aware of problematic prescribing situations and would come to avoid prescribing practices that were sub-optimal,.or which lead to risk for adverse effects. It was hypothesized, therefore, that the cognitive services demonstration in Washington should be effective in reducing the rate of screen failures in comparison to pharmacies that were not actively engaged in providing cognitive services.

### Evaluation Approach

The major objective of the present study was to examine the impact of the cognitive services demonstration on the occurrence of screen failures associated with the eight pharmaco-therapeutic drug classes that are covered in the PSU Screener. The expectation was that pharmacists in the demonstration would be more likely than control pharmacists to detect prescription problems and to consult with prescribers regarding therapeutic alternatives. Over time, as pharmacists had increased contact with prescribers and had opportunities to impact on their prescribing practices, the incidence of screen failures was therefore expected to decline.

To address these issues, a multivariate analytic design was employed. For the screen failure outcome, defined as any screen failure during the follow-up period within a given therapeutic class, the residual effect of the demonstration was examined after adjustment for known risk factors. Multivariate regression procedures were used to control for patient demographic variables and drug utilization which could affect a given patient's likelihood of new screen failures. All control variables were measured during a baseline period prior to the onset of the demonstration so as to be independent of any demonstration effects.

### Data

The data used in this study was extracted from Medicaid drug claims and Medicare files for patients in the state of Washington in 1993 (to provide a baseline) and 1994-1995. Other records and data streams generated by the demonstration were also incorporated into the data set.

### *Data Base and File Structure*

Pharmacies were recruited into the demonstration in three distinct waves (February, April, and September of 1994). This study assumes that a patient's exposure to the demonstration (and consequently the patient's follow-up period) began when a patient first filled a prescription at a pharmacy participating in the demonstration. In the majority of cases, the follow-up period then continued until the demonstration ended in September 1995.

All persons in the study were ascertained to be eligible for Medicaid for the full period of the demonstration. In certain cases, if the patient lost eligibility for Medicaid, entered a nursing home, or entered an HMO, the follow-up period ended sooner because the prescription drug claims data would be incomplete. A patient's death was also deemed to terminate the follow-up period.

Each patient's baseline periods might also differ in length. The baseline period was determined to end the day before the patient's first contact with a demonstration pharmacy and to reach back twelve months or until the patient first became eligible for the Medicaid program, left an HMO, or left a nursing home and became resident in the community. Again, the latter three restrictions are imposed in order to ensure that the drug claims records are complete during the baseline period. These windows were constructed in order to capture distinct time periods during which the demonstration could and could not have influenced screen failure rates.

### *Study Populations*

The main population for this study included Washington residents who were eligible for Medicaid and who filled prescriptions at pharmacies participating in the cognitive services demonstration. For the analyses presented in this chapter, additional requirements were that participants must: 1) have three or more months of baseline data; 2) have an adequate follow-up period; 3) not be enrolled in a managed care program; and 4) not reside in a nursing home.

The final population of eligible study participants included individuals both younger than and older than age 65. However, for the present study, the elderly represent a population of particular interest because on average the elderly experience poorer health and receive more prescription drugs than younger individuals. These factors, along with aging-related alterations in pharmacokinetics, place the elderly at much greater risk for numerous adverse drug reactions. It was expected that the elderly should represent a subgroup to which the cognitive services would most likely be provided and therefore a subgroup that would most likely experience a decline in screen failures. Moreover, even among the elderly there are subgroups whose drug therapy is highly likely to present problems: elderly receiving prescription from multiple prescribers, and elderly receiving a large number of prescription within a relatively short time period. Lastly, it was thought that diabetics, because of the nature of their disease condition and/or drug therapy were expected to result in greater scrutiny by pharmacists proactively providing cognitive services.

Exhibit 4.6.1 summarizes the selection of the final study population and subsamples from the original pool of Medicaid eligible persons who filled prescriptions at any demonstration pharmacy between 1993 and 1995 (N=254,032). A total of 106,706 individuals were eliminated due to incomplete baseline or follow-up data. An additional 55,932 individuals were eliminated due to the timing of their

**Exhibit 4.6.1**

**HCFA DUR Demonstration Evaluation**
**Derivation of Washington Project C.A.R.E. Screen Failure Analysis Samples**

| | |
|---|---|
| Medicaid eligible, filled prescriptions at demonstration pharmacy | 254,032 |
| *Less:* | |
| Individuals with insufficient data (no follow-up period, baseline less than three months) | (106,706) |
| Not exposed due to timing of first prescription | (51,812) |
| Filled prescription in both arms of study | (4,120) |
| Net usable study population | 91,394 |
| Dually eligible (Medicaid/Medicare) | 16,904 |
| Non-elderly (< 65 at end of study) | 8,080 |
| Became 65 during study (omitted from sample) | (1,015) |
| Elderly (65+ at beginning of study, 9/30/95) | 7,809 |
| *Subpopulations of elderly eligibles:* | |
| Elderly diabetic | 1,680 |
| Elderly with 3+ prescribers in any baseline month | 1,749 |
| Elderly with 10+ prescriptions in any baseline month | 952 |

Source: Washington Project C.A.R.E. Medicaid/Medicare Files

first prescription or because they utilized pharmacies in both arms (control and demonstration) of the study.

The final study sample of 16,904 dually-eligible recipients included similar numbers of individuals under 65 (N=8,080) and those 65 or older (N=7,809). The elderly population (N=8,080) was further subset into groups especially likely to benefit from the cognitive services demonstration. These included:

1) Elderly diabetics – individuals who had claims for insulin or oral hypoglycemic agents at any point during 1993-1995 (N=1,680).

2) Elderly with three or more prescribers during any baseline month (N=1,749)

3) Eligible elderly with 10 or more distinct drugs (as represented by 9-digit NDC's) during any single baseline month (N=952).

### Dependent Variables

All of the outcomes selected for this portion of the study involved the occurrence of screen failures during the follow-up period. Downstream screen failure measures included indicators both of whether or not a screen failure occurred, and, for those with any screen failures, how many failures were generated during the follow-up period.

The DUR criteria screened for this study were originally developed by expert panels convened by the University of Maryland and the Philadelphia College of Pharmacy and Science (UM/PCPS) under a cooperative agreement with HCFA (Knapp and Erwin, 1992) and which are generated by the PSU Screener. Additional details regarding the PSU Screener and the specific DUR criteria used in this study have been provided in Chapters 2 and 3 of this report.

The screening process involved a review of all Medicaid prescriptions within the following drug categories: ACE inhibitors, antidepressants, antipsychotics, benzodiazepines, calcium channel blockers, digoxin, $H_2$-receptor antagonists, and non-steroidal anti-inflammatory agents. The types of drug problems screened included four prescribing problems — dosage, drug/drug interactions, duration of therapy, and therapeutic duplication.

The study employed both broadly defined measures, such as any screen failure among the group of specific screening criteria organized by therapeutic class, and also more narrowly-defined measures involving the specific screens within the eight therapeutic class. In order to examine the impact of the intervention on screen failure rates, analyses were first conducted for each of the eight drug classes represented in the PSU screener criteria and for an aggregate variable representing any screen failure over the total 61 screens contained in all eight therapeutic classes. Additional analyses were conducted on a subset of 30 of the 61specific screens that are generated by the PSU Screener. These 30 screens were selected based on an observed annual prevalence greater than 5% in any of the four states for which screen prevalence rates were computed.[22]

### Independent Variables

In modeling the effects of the demonstration on subsequent screen failures, measures were first developed to characterize demonstration exposure, then, based on the available data, additional covariates were identified which were hypothesized to impact on screen failure rates independently of the effects of the demonstration. Independent variables selected for these analyses are listed in Exhibit 4.6.2 and are briefly detailed below.

#### *Exposure to the Intervention*

Although this demonstration was designed as a random experiment, randomization occurred at the level of the pharmacy cluster while the impact of the intervention is measured at the level of the

---

[22] Three screens were omitted from this rule because of low denominators: Nimodipine duration, Ketrolac with probenecid, and Mefenamic acid or phenybutazone duration.

**Exhibit 4.6.2**

**HCFA DUR Demonstration Evaluation**
**Independent Variables: Washington Cognitive Services Downstream Screen Failure Analysis**

| Variable | Definition |
|---|---|
| *Demonstration exposure* | |
| Treatment (0,1) | Demonstration status of individual: treatment if all prescriptions during study period filled at treatment pharmacies, control if all filled at control pharmacies |
| *Other independent variables* | |
| Age | Nine qualitative variables: 20-29, 30-39, 40-49, 50-59, 60-64, 65-69, 70-79, 80-84, 85+ |
| Sex | Two qualitative variables: male, female |
| Race | Three qualitative variables: white, black, other |
| Eligibility | Medicaid eligibility - blind or disabled |
| DCG/HCC | Predicted annual Medicare payments (proxy measure of medical need) |
| Baseline screen failures | Pharmacy cluster average of screen failures during the baseline period |
| Baseline drugs | Maximum number of distinct drugs in a single month during the baseline period; natural logarithm of value entered into regression models |
| Length of baseline period | Measured in months; natural logarithm of value entered into regression models |
| Length of follow-up period | Measured in months; natural logarithm of value entered into regression models |
| Wave 1 | Wave 1 of pharmacy entrance (for pharmacy filling individual's first prescription) into demonstration (three waves in Washington) |
| Wave 2 | Wave 2 of pharmacy entrance (for pharmacy filling individual's first prescription) into demonstration |

Source: Washington Project C.A.R.E. Medicare/Medicaid Files

patient.[23] This presented a significant methodological problem. Individuals had the potential to fill prescriptions at different pharmacies and, potentially, to fill prescriptions at a mix of intervention, control, and non-demonstration pharmacies. It could be expected that the effect of the demonstration would be diluted for such individuals depending on the relative proportions of prescriptions filled at the different pharmacy types. To resolve this problem, this study adopted the following approach. Individuals were considered members of the intervention group if they ever had a Medicaid-paid drug claim at an intervention pharmacy and never had such a claim at a control pharmacy. Similarly, they were considered members of the control group if they ever had a Medicaid-paid drug claim at a control pharmacy and never had such a claim at an intervention pharmacy. Individuals who went to both intervention and control pharmacies were excluded from the study regardless of the number of filled prescription at either intervention or control pharmacies.[24]

---

[23] Pharmacy clusters, built around common prescribers, were created and randomized to minimize the influence that intervention pharmacies might exert through common prescribers on control patients. In Washington there were 38 pharmacy clusters ranging in size from 1 person to 1200 people.

*Other Covariates*

***General Predisposing and Need Characteristics*** – The basic strategy for the analyses reported here was to examine the effect of the treatment variable using regression models to statistically controlled for demographic factors and other measures thought to be related to the risk of invoking a screen failure. For all dependent variables, the regression models included variables for the patient's race, sex, and age. Additionally, a dichotomous variable was created to indicate whether or not a person's Medicaid eligibility was associated with being blind or disabled. Another variable was created to control for medical need, independent of the need that was influenced by this demonstration. The principal measure of need was based on the Diagnostic Cost Group (DCG) and Hierarchical Coexisting conditions (HCC) models and software developed by Ellis et al. (1996). The DCG/HCC model uses information on a person's age, sex, diagnoses, life-sustaining medical procedures, and selected hospitalizations during the baseline period and uses that information to predict Medicare payments for that person in the next year. As an additional measure of need, regressions controlled for the maximum number of distinct drugs that each individual received during a single month of the baseline period. These factors could affect screen failure rates indirectly as proxies for health status. They may also affect the probability of having access to, or using, health care services. Thus these measures serve to control for general predisposing and need characteristics.

***Rate of Screen Failure in Pharmacy Cluster*** – The pharmacy cluster, the unit of randomization in this analysis, effectively defined medical communities. Potentially, there could be unmeasured differences in demographic and behavioral characteristics among patients in different pharmacy clusters that influenced the need for care and the predisposition to seek it. Similarly, there could be geographic variation associated with cluster in access to care and in its practice, such as variations in coding of conditions and diagnoses and in treatment protocols. If these differences were associated with measured clinical outcomes, then controlling for them would improve the precision of estimated intervention effects. Thus, each model included the average screen failure rate during the baseline period for each pharmacy cluster as an additional regressor. This variable served as a proxy for the net effect of geographic factors enhancing or reducing access to care and for geographic variation in the practice of care, such as variations in prescribing patterns for specific conditions.

***Length of Baseline and Follow-up Periods*** – It would be expected that patients with longer follow-up periods would be more likely to experience one or more screen failures than those with shorter follow-up periods. Therefore, the length of the follow-up period was included as a covariate in all models. While the length of the baseline period does not directly affect patients' probabilities of generating screen failures, it is likely that the assessment of drug utilization and other risk factors may be measured with less precision for patients with short baseline periods relative to patients with full baseline periods. Thus, controlling for other measured covariates, the probability of screen failures might vary according to the length of the baseline period. For all models, the lengths of the follow-up and baseline period were entered as the natural log of the number of months in each period.

---

[24] For this purpose, the day that the patient first went to the demonstration pharmacy was included in the follow-up period.

***Pharmacy Wave*** – Pharmacies entered the cognitive services demonstration in three waves. The first wave of 160 pharmacies enrolled in February 1994. A second wave (N = 33) and a third wave (N = 7) of pharmacies entered in April and September 1994. It is possible that pharmacies entering the demonstration later learned from the experiences of prior entrants and provided more effective cognitive services. Additionally, the rate of cognitive service events per 100 dispensed prescriptions varied widely over time in the treatment group. For these reasons, indicator variables were included in the models that account for the wave of the first demonstration pharmacy where the individual filled a prescription. The reference category in the models was wave one. The basic set of independent variables that were used in every model (as described above) are summarized below in Exhibit 4.6.2.

### Models and Statistical Methods

Because the dependent variables denoting any screen failure were categorical and binary, probit regressions were used. The regression specification for the examination of treatment effects in this study was:

$$y^{c}_{ijt_f}=\alpha+\beta_1 tr_{ij}+X_{it_b}\beta_2+W^{c}_{it_b}\beta_3+Z_i\beta_4+\beta_5 y^{c}_{jt_b}+\nu_{ij}$$

with:

| | |
|---|---|
| $y^c_{ijtf}$ | indicator that patient i associated with pharmacy cluster j experienced screen failure c during the follow-up period (tf) |
| $tr_{ij}$ | indicator that patient i filled a prescription at treatment pharmacies and not at control pharmacies during the demonstration period |
| $X_{it}$ | general predisposing (socio-demographic) and need characteristics of patient i determined during the baseline period |
| $W^c_{itb}$ | patient i's risk factors for screen failure c determined during the baseline period |
| $Z_i$ | length of follow-up and baseline windows for patient i |
| $y^c_{jtb}$ | rate of screen failure c in pharmacy cluster j during the baseline period |
| $\nu_{ij}$ | error |

### Findings

#### *Descriptive Statistics*

Exhibit 4.6.3 below displays the means of the independent variables included in the regressions, for all five analysis populations used in the study. Some differences between the elderly and non-elderly groups primarily reflect programmatic rules. For example, the non-elderly group is mainly comprised of blind or disabled participants (96%), compared to 13% in the elderly group. Elderly recipients in general had higher expected medical expenses and took more baseline drugs than non-elderly. Relative to the elderly population, a much higher percentage of the non-elderly were male (50% versus

Exhibit 4.6.3

### HCFA DUR Demonstration Evaluation
### Washington Project C.A.R.E Screen Failure Analysis Sample Characteristics

| | Population | | | | |
|---|---|---|---|---|---|
| Measure | Non-elderly (N=8,080) | All elderly (N=7,809) | Diabetic elderly (N=1,680) | Elderly with 3+ pre-scribers (N=1,749) | Elderly with 10+ Rx's (N =952) |
| Age (30-39) | 0.32 | | | | |
| Age (40-49) | 0.28 | | | | |
| Age (50-59) | 0.20 | | | | |
| Age (60-64) | 0.04 | | | | |
| Age (70-79) | | 0.44 | 0.45 | 0.44 | 0.46 |
| Age (80-84) | | 0.14 | 0.12 | 0.12 | 0.11 |
| Age (85+) | | 0.12 | 0.07 | 0.07 | 0.06 |
| Race (black) | 0.06 | 0.05 | 0.08 | 0.06 | 0.05 |
| Race (other) | 0.08 | 0.24 | 0.26 | 0.17 | 0.16 |
| Sex (male) | 0.50 | 0.26 | 0.25 | 0.23 | 0.21 |
| Blind or disabled | 0.96 | 0.13 | 0.15 | 0.15 | 0.15 |
| Wave 2 | 0.12 | 0.17 | 0.17 | 0.16 | 0.14 |
| Wave 3 | 0.04 | 0.04 | 0.05 | 0.04 | 0.06 |
| Treatment group | 0.61 | 0.59 | 0.60 | 0.59 | 0.59 |
| DCG/HCC | $8,281 | $8,440 | $8,852 | $8,753 | $9,033 |
| Baseline drugs | 2.5 | 2.6 | 5.0 | 7.74 | 12.2 |
| Baseline length (months) | 10.1 | 10.4 | 10.5 | 11.1 | 10.9 |
| Follow-up length (months) | 10.3 | 10.1 | 10.6 | 10.8 | 9.9 |

Source: Washington Project C.A.R.E. Medicare/Medicaid Files

26%). The three elderly subpopulations differ from the total elderly population in several respects: not surprisingly, diabetic or high-utilization elderly have higher expected medical costs and use more drugs than the overall elderly group. The five populations do not differ greatly in the proportion of individuals visiting pharmacies in the treatment groups or pharmacies enrolling in the demonstration during Wave 2 or Wave 3.

### *Treatment Effects*

It was hypothesized that, relative to individuals visiting control pharmacies, individuals who visited pharmacies providing cognitive services would be less likely to experience screen failures during the follow-up period. If the feedback to patients and their prescribers did in fact significantly decrease subsequent prescribing problems likely to generate DUR screen alerts, then the models should indicate that, after adjusting for baseline demographic and risk factors and for factors related to potential design effects, the treatment group should have had lower rates of screen failures. If this were the case, the probit coefficients should be negative and significant. Similarly, among groups of individuals who experienced any screen failure during the follow-up period, the OLS coefficients obtained from predicting total number of screen failures from the treatment variable should also be negative and significant.

***Effects with the eight therapeutic classes*** – Exhibits 4.6.4 provides the coefficient signs and their respective statistical levels of significance for the demonstration treatment effects within the eight therapeutic classes defined by the PSU Screener. Results for all probit models are presented; results for OLS models are presented for those outcomes for which OLS models were run. (OLS models were only performed for subpopulations or drug categories in which the total number of persons experiencing screen failure was 100 or greater.)

Examination of the estimation results summarized in Exhibit 4.6.4 suggests that the Washington Project C.A.R.E. system was not greatly effective in preventing and reducing screen failures during the follow-up period. For the probit models, which evaluated the effect of treatment on any screen failure while controlling for other variables, 14 out of the 45 estimated models (31%) yielded treatment coefficients that were negative in sign. Only one of these coefficients, however, was statistically significant at a level of 0.05; elderly with 3 or more prescribers, for the screen of any H2RA failure (p=.047). In addition to these findings, three probit coefficients for the treatment effect were positive and significant. This effect occurred for failures involving use of ACE Inhibitors for three subpopulations: elderly (p. = .005), elderly with three or more prescribers (p = .027), and elderly with 10+ prescription drugs (p = .024).

In addition to the probit models, OLS regression models were also performed which examined the association between the independent variables and the number of screen failures generated during the follow-up period. In order to ensure the minimal sample size necessary for model estimation, these OLS regressions were only performed within sub-populations and drug classes in which the total number of individuals experiencing one or more screen failures exceeded 100. The available OLS results, which are also presented in Exhibit 4.6.4, do not suggest any clear effects of the treatment exposure on level of screen failures among individuals experiencing any screen failure. Only 12 of the 31 OLS regression coefficients for the treatment effect variables were negative in sign. None of the of the treatment coefficients were significant. One positive OLS coefficient was statistically significant: elderly using 10+ prescription drugs, H2RA's (p=.007).

The positive associations of treatment with screen failures is unexpected. However, it is unlikely that these positive effects of treatment on screen failure level are indicative of any systematic response to the intervention. In order for systematic positive associations to have occurred, either pharmacists

**Exhibit 4.6.4**

**HCFA Demonstration Evaluation**
**Regression Adjusted Estimates of Total Utilization Effects: All Analysis Populations**

| Therapeutic Category | Non-elderly | All Elderly | Elderly diabetics | Elderly with 3+ prescribers | Elderly with 10+ drugs |
|---|---|---|---|---|---|
| ACE Inhibitors | | | | | |
| Any failure | (-) | (+)** | (+) | (+)* | (+)* |
| Number of failures | (+) | (-) | ---- | ---- | ---- |
| Antidepressants | | | | | |
| Any failure | (+) | (+) | (-) | (-) | (+) |
| Number of failures | (+) | (-) | ---- | (+) | (-) |
| Antipsychotics | | | | | |
| Any failure | (+) | (+) | (-) | (+) | (+) |
| Number of failures | (+) | (-) | ---- | ---- | ---- |
| Benzodiazepines | | | | | |
| Any failure | (-) | (-) | (+) | (-) | (+) |
| Number of failures | (-) | (+) | (+) | ---- | (+) |
| Calcium channel blockers | | | | | |
| Any failure | (+) | (+) | (+) | (+) | (-) |
| Number of failures | (-) | (+) | (-) | (+) | ---- |
| Digoxin | | | | | |
| Any failure | (-) | (+) | (+) | (+) | (+) |
| Number of failures | ---- | (-) | ---- | ---- | ---- |
| H2RA's | | | | | |
| Any failure | (+) | (-) | (-) | (-)** | (-) |
| Number of failures | (-) | (+) | (-) | (-) | (+)** |
| NSAID's | | | | | |
| Any failure | (+) | (+) | (-) | (+) | (+) |
| Number of failures | (+) | (+) | (+) | (+) | (+) |
| All Drug Classes | | | | | |
| Any failure | (+) | (-) | (+) | (-) | (+) |
| Number of failures | (+) | (+) | (+) | (-) | (+) |

| | |
|---|---|
| Source: | Washington Project C.A.R.E. Medicare/Medicaid Files |
| ---- | OLS not performed (N < 100) |
| * | The difference between the arms is statistically significant at the 0.10 level |
| ** | The difference between the arms is statistically significant at the 0.05 level |
| *** | The difference between the arms is statistically significant at the 0.01 level |

or prescribers would have to have responded unfavorably to the cognitive services intervention and in a manner that increased subsequent screen failure rates. It seems unlikely, even if prescribers disagreed with, or responded negatively to, the implementation of cognitive services that they would actively increase their rate of prescribing of regimens that would fail the screening criteria used in the PSU Screener.

*Effects within specific screening criteria* – In addition to examining the treatment effects by broad therapeutic class, examination of the treatment effects for specific criteria may also yield useful information regarding the effectiveness of the cognitive services intervention in reducing specific screen criteria. Exhibit 4.6.5 provides the probit coefficient signs and their respective statistical levels of significance for the demonstration treatment effects for the 30 specific screen criteria with prevalence rates high enough to warrant selected for further analyses.[25]

Within specific criteria, only three probit coefficients for the treatment effect were statistically significant at the conventional cutoff of $p < .05$ and of these only 1 (benzodiazepines with anxiolytics) was negative. Of the 146 coefficients that were estimated, 64 or 44% were negative. It should be noted, however, that the sample sizes for many of the individual criteria were very small and statistical power to detect treatment effects was limited at the level of individual screening criteria.

Overall, the results shown in Exhibits 4.6.4 and 4.6.5 indicate little treatment effect on the occurrence of screen failures within the eight drug classes screened and for the total across all drug classes. No discernable patterns involving specific patient sub-populations or therapeutic classes are apparent. These findings suggest that the cognitive services intervention was relatively ineffective in changing the prescribing patterns that might lead to the kinds of screen failures that were included in the PSU Screener.

### *Other Covariates*

Race and gender generally were not strong predictors of screen failure in the populations studied here, although there were a few exceptions. For example, in the non-elderly, black and other race were associated with lower rates of screen failures for benzodiazepines ($p < 0.01$). Among elderly, other race was also associated with lower rates of screen failures for antidepressants ($p < .001$), antipsychotics, ($p <.001$), and NSAIDS ($p < .01$). For the non-elderly, male sex was strongly associated with lower rates of screen failures within the broad therapeutic classes of antidepressants ($p < 0.001$), benzodiazepines, ($p <.01$), and NSAID's ($p < 0.001$). However, among elderly, other race was associated with higher rates of screen failures for Ace inhibitors ($p < .01$). Not surprisingly, for most therapeutic classes, measures of medical need (such as predicted Medicare expenditures) and drug utilization (such as maximum number of drugs used during a baseline month) were strongly associated with higher rates of screen failure. Further, for most classes, length of follow-up was a strong positive predictor of risk for screen failures.

---

[25] Despite requiring the cutoff for the prevalence rate for screen failures to be higher than 5%, there were a number of instances where the probit or OLS analyses could not be conducted due to low occurrences of the screen failures.

**Exhibit 4.6.5**

### HCFA DUR Demonstration Evaluation
### Regression Adjusted Estimates of Treatment Effects: 30 Specific Criteria

| Screen Class and Name | Non-elderly | All elderly | Elderly diabetic | Elderly 3+ pre-scribers | Elderly 10+ Rx's |
|---|---|---|---|---|---|
| ACE Inhibitors | | | | | |
| Multiple ACE inhibitors | (-) | (+)* | (+)* | (+) | (+)* |
| Antidepressants | | | | | |
| Adult dosage | (+) | (+) | (+) | (+) | (+) |
| Multiple antidepressants | (+) | (+) | (+) | (-) | (-) |
| Tricyclic antidepressant or maprotiline with clonidine | (+) | (-) | (-) | (-) | (+) |
| Antipsychotics | | | | | |
| Adult and pediatric dosage | (+) | (-) | (-) | (+) | (-) |
| Antipsychotic with antiparkinsonian | (-) | (+) | (+) | (+) | (+) |
| Multiple antipsychotics | (+)** | (+) | (+) | (-) | (-)* |
| Benzodiazepines | | | | | |
| Adult dosage | (+) | (+) | (+) | (-) | (-) |
| Benzodiazepines, duration | (-) | (+) | (-) | (-) | (-) |
| Benzodiazepine with antihistamine | (-) | (+) | (+)** | (-) | (-) |
| Benzodiazepine with anxiolytic | (-) | (-) | (+) | (-) | (-) |
| Benzodiazepine with sedative/hypnotic | (-)*** | (-) | (-)* | (-) | (+) |
| Estazepam, flurazepam, temazepam, or triazolam, duration | (+) | (-) | (-) | (-) | (-) |
| Long half-life benzodiazepine over age 65 | na | (+) | (+) | (+) | (-) |
| Multiple benzodiazepines | (+) | (+) | (+) | (-) | (-) |
| Calcium Channel Blockers | | | | | |
| Adult dosage | (-) | (+) | (+) | (+) | (-) |
| Multiple calcium channel blockers | (-) | (+) | (+) | (+) | (+) |
| Verapamil with beta blocker | (+) | (-) | (-) | (+) | (+) |
| Digoxin | | | | | |
| Adult dosage | (-) | (+) | (-) | (+) | (+) |
| $H_2$RA's | | | | | |
| Adult dosage | (+) | (-) | (-) | (+) | (-) |

**Exhibit 4.6.5**

**HCFA DUR Demonstration Evaluation**
**Regression Adjusted Estimates of Treatment Effects: 30 Specific Criteria**

| Screen Class and Name | Non-elderly | All elderly | Elderly diabetic | Elderly 3+ pre-scribers | Elderly 10+ Rx's |
|---|---|---|---|---|---|
| Cimetidine with phenytoin, methylxanthines, or warfarin | (-) | (+) | (+) | (+) | (+) |
| $H_2$RA acute duration | (-) | (-) | (-) | (-) | (+) |
| $H_2$RA and sucralfate on same day | (+) | (-) | (-) | (+) | (+) |
| $H_2$RA GERD duration | (-) | (-)* | (-) | (-) | (-) |
| $H_2$RA with misoprostol, omeprazole, or sucralfate | (+) | (+) | (+) | (-) | (-) |
| Multiple $H_2$RA's | (-) | (+) | (-) | (-) | (-) |
| NSAID's | | | | | |
| Adult dosage | (+) | (+) | (-) | (+) | (-)* |
| Mefenamic acid or phenylbutazone duration | (-) | na | na | (+) | na |
| Multiple NSAID's | (-) | (+) | (-) | (+) | (+) |
| Salicylates with oral antidiabetic | (+) | (-) | (+) | (-) | (-) |

Source: Washington Project C.A.R.E. Medicare/Medicaid Files
na The number of screen failures was too small for probit convergence to occur
* The difference between the arms is statistically significant at the 0.10 level
** The difference between the arms is statistically significant at the 0.05 level
*** The difference between the arms is statistically significant at the 0.01 level

### *Pharmacy Cluster Effects*

To account for the fact that clusters of pharmacies rather than individuals were randomized to the treatment and control groups, each model included as a covariate the average rate of baseline screen failure in the individual's pharmacy cluster. It was expected that this variable would help control for unobservable, cluster level differences in factors that affected the frequency of screen failures. Across the five populations and eight drug classes screened, these variables were consistently and very strongly associated with screen failure.

### Comment

The Washington cognitive services intervention was designed to offer appropriate financial compensation to assist pharmacist to provide drug therapy counseling to patients and physicians. The goals for the payment for cognitive service was to reduce drug therapy risks for patients, and to improve the quality of prescribing practices by prescribers. This chapter focused on the impact of the intervention on the incidence of screen failures. The analysis controlled for treatment and comparison group characteristics that may have differed at baseline, as well as the cluster-based sampling design used in the demonstration. Based on the regression results discussed above, there is little evidence that the cognitive services interventions prevented subsequent screen failures.

There are a number of reasons that may explain why no significant effect this effect was found. Because the intervention was not specifically focused on the criteria contained in the PSU Screener, it may be that the screening criteria used here were not appropriate for evaluating the impact of the intervention. While unlikely, it may be that most, if not all, of the screening criteria used here were not specific enough to capture the type of problems that were addressed by the demonstration pharmacists.

The attitudes and behaviors of the prescribing physician are a second critical component influencing the outcome of intervention. In general, reduction in therapeutic problems and the incidence of screen failures will be limited primarily by the willingness and ability of physicians to modify their prescribing patterns. Without additional data indicating what specific actions were taken by pharmacists and prescribers in response to the cognitive services, it is difficult to draw any strong conclusions regarding which screening criteria might have been constructed and implemented in the PSU Screener to best evaluate the intervention.

It is also likely that limitations in the study design complicated the measurement and detection of existing treatment effects, particularly if the treatment effects were very small. Generally, a cluster sampling approach reduces the power of detecting differences between the treatment and control groups. In this case, however, the intracluster correlation is likely to be low, and many of the clusters were very small; cluster based sampling therefore does not seriously reduce the power of the study design. More important is the fact that, if the treatment effect is expected to be very small then the study had limited power to detect a differences. To capture small differences between treatment and control groups, larger sample sizes would be needed to receive the intervention. Although these sample size constraints greatly limited the ability of the study to detect significant effects, the findings nevertheless suggest that the cognitive services demonstration did not lead to large declines in screen failure rates, for the drug therapy screens that were examined.

## 4.7 Outcomes Analyses – Methods and Data

This report presents findings on the impact of Project C.A.R.E. on adverse outcomes related to use of four classes of prescription drugs — NSAIDS, antidepressants, cardiovascular drugs, and benzodiazepines. These four studies are presented in separate chapters, and each study used methods that were designed to appropriately address the specific clinical issues that were of interest. Nevertheless, all of the outcomes studies were based on the same data and variable creation methodology. Furthermore, to a large extent, the statistical models used in the four studies were very similar. The objective of this chapter is to present an overview of the data, general evaluation approach and variable creation methods that were used across all of the outcomes studies. Each outcomes study chapter summarizes this general approach and also describes methods and variables used that were specific to the study.

## General Evaluation Approach

The objective of each outcomes study was to explore the relationship between payment for cognitive services and the rate of adverse events related to use of a particular set of prescription drugs. It was hypothesized that payment for cognitive services would increase the number of cognitive service interventions provided. This increase, in turn, was expected to change prescriber and/or patient behavior, and make prescriptions safer and more effective. The rate of adverse outcomes related to prescribing was therefore expected to decrease, reflecting improvements in prescribing and patient compliance.

Project C.A.R.E. randomized pharmacy clusters to treatment and control groups. The treatment group received a participation fee as well as reimbursement for cognitive services provided while the control group received a participation fee only. In the analysis, individuals were considered to have received treatment if they filled a Medicaid paid prescription at a treatment pharmacy at least once during the study period. Individuals were considered to be members of the control group if they filled a Medicaid paid prescription at a control pharmacy at least once during the study period. To avoid problems with crossover, individuals who visited both treatment and control pharmacies were dropped from the analysis.

Each of the outcomes studies essentially compared the treatment and control groups' rates of clinical outcomes. Because clusters were randomized instead of individuals, there were some demographic differences between the treatment and control groups. For this reason, multiple regression was used to assess the effect of the treatment on individuals' outcomes. The models controlled for demographic, disease, and drug factors predisposing a given patient to the outcome under study. All control variables were measured during a baseline period before the demonstration began so as to be independent of any demonstration effects. The treatment was measured using an indicator variable. Because all dependent variables were binary, probit regressions were used.

Thus, the regression specification was:

$$y^{c}_{ijt_f}=\alpha+\beta_1 tr_{ij}+X_{it_b}\beta_2+W^{c}_{it_b}\beta_3+Z_i\beta_4+\beta_5 y^{c}_{jt_b}+\nu_{ij}$$

with:

| | |
|---|---|
| $y^{c}_{ijtf}$ | indicator that patient i associated with pharmacy cluster j experienced clinical outcome c during the follow-up period (tf) |
| $tr_{ij}$ | indicator that patient i filled a prescription at treatment pharmacies and not at control pharmacies during the demonstration period |
| $X_{it}$ | general predisposing (socio-demographic) and need characteristics of patient i determined during the baseline period |
| $W^{c}_{itb}$ | patient i's risk factors for outcome c determined during the baseline period |
| $Z_i$ | length of follow-up and baseline windows for patient i |
| $y^{c}_{jtb}$ | rate of clinical outcome c in pharmacy cluster j during the baseline period |
| $\nu_{ij}$ | error |

### *Covariates*

Each outcomes study used as covariates a distinct set of variables that were expected to influence the outcomes under investigation. These variables are described in the outcomes study chapters that follow this chapter. A basic set of covariates, however, was used in every model across the four studies. In addition to the treatment dummy variable, this set included race, sex, age, length of baseline period, length of follow-up period, and the wave when the pharmacy entered the demonstration.

### *Correcton for Cluster Sampling*

The fact that pharmacy clusters rather than individual patients were randomized required that the standard regression framework be adjusted. If indeed there were factors correlated with cluster that affected outcomes and if, despite the controls included in the models, these factors were not fully measured, then, because the number of clusters was small (38) relative to the number of patients, these differences might not average out between intervention and control groups, leading to misleading estimates of intervention effects. Specifically, traditionally-estimated standard errors would be smaller than the true standard errors.

The first line of defense against this problem was the inclusion of a full set of covariates, including a variable that measured the cluster average of the dependent variable. This cluster average was expected to capture factors that vary by cluster that are important in the determination of clinical outcomes. In addition, a cluster adjusted version of White's method was used to estimate the variance-covariance matrices of the models. This method relaxes the assumption of independence within clusters, and is considered to be a robust variance estimator in cases of cluster-based sampling. An alternative approach would have been to bootstrap the coefficients of each model and derive nonparametric standard errors for the coefficients.[26] In two of the outcomes studies (antidepressant and cardiovascular drugs), this alternative method was tested; each anti-depressant and cardiac model was bootstrapped and standard errors were computed. The results, which are not presented in this report, were very similar to the results derived using the cluster-adjusted version of White's method.[27]

## Data and Variable Creation

The analysis of selected health outcomes among users of demonstration pharmacies is based on reviews of data derived from Medicaid and Medicare administrative sources. Administrative data is collected for the purposes of program management and to enable and record the payment of medical providers for services and supplies. The data is not collected for health research purposes, however there are a number of data fields that contain patient clinical and demographic information that can be used for the analysis of health outcomes. The data used in the analyses are demographic and program participation information from enrollment data and drug use, diagnosis and procedure information from medical claims billing records. The data available for the analyses was complete state Medicaid and Medicare data for services delivered in the period 1993-1996.

---

[26] The bootstrapping method involves pulling repeated samples (with replacement), estimating coefficients repeatedly (100 times in this case), and then deriving standard errors from the distributions of estimated coefficients. This method is often used in cases where we have no formulas to compute standard errors or when the assumptions we need to make to compute standard errors do not make sense (i.e. in this case, we could not assume that error terms were independently distributed because of inter-cluster correlation).

[27] Two other methods were considered to account for the effects of clustered sampling. The random effects probit was considered, but the large size of the clusters made this model very difficult if not impossible to implement computationally. Generalized Estimating Equations (GEE) methods could have been used, but this approach would have made estimation more complex and the advantages of this approach were uncertain.

### *Enrollment Data*

The populations potentially affected by demonstration influences are Medicaid recipients eligible for pharmacy benefits. A significant portion of the Medicaid population are also entitled to benefits under the Medicare program. Enrollment data from both programs contains person-level information that details by date enrollment status. Included in these records are data describing personal demographic characteristics, managed care affiliation and the specific categories under which eligibility has been granted. From an epidemiological perspective the enrollment data is the source of all population denominator information. Population studies that focus on disease prevalence and the frequency of acute events require program eligibility data to define and model study groups. The models require the demographic information but also need to account for individuals who are not eligible over the study period and whose claims history is incomplete.

Most of the outcome analyses focused on the community dwelling elderly. This population is generally covered under Medicaid for pharmacy benefits and Medicare for institutional and physician services. Medicaid recipients who are not entitled to Medicare benefits have complete coverage under the Medicaid program for all medical products and services. To fully model the population the Medicaid and Medicare enrollment data were linked using common identifiers. The cross-program linkages were validated by ensuring matches based on coded identifiers as well as personal characteristics such as birthdate and gender. The integration of the enrollment data allowed for the monthly determination of eligibility status in both programs. The preparation of a joined Medicaid and Medicare enrollment file entailed the assignment of unique study identifiers to all records pertaining to an individual. Enrollment data is used in the analyses to identify eligible populations, determine baseline and follow-up observation windows and to assign values to the demographic covariates.

### *Pharmacy Claims*

Medicaid pharmacy claims are used in the outcome analyses to both identify the populations directly affected by the demonstration and to document drugs used by the patients in the baseline and follow-up periods. A pharmacy claim record reliably includes the following data elements: patient identifier, pharmacy identifier, date of dispense, drug entity and quantity dispensed. Prescription duration information is reported in pharmacy claim records but it is of variable quality. The outcome analyses uses drug claims data to identify Medicaid recipients who used demonstration pharmacies. Drug claims are also used to identify population risk groups based on pre-demonstration exposures to targeted therapeutic entities. Drug claims data served as a source for proxy information on diagnostic status when the presence of a therapy was a strong indicator of disease.

### *Physician Service Claims*

Medicaid and Medicare physician claims are used in the outcome analyses as source of diagnostic information and to detect the occurrence of targeted medical procedures. Physician claim records reliably report the following data elements: patient identifier, provider identifier, dates of service (at least to the month level), medical procedures performed and diagnoses. Diagnostic and procedure information are used in the analyses to determine outcome status in the follow-up period and for the formulation of baseline clinical covariate values. Two of the outcome analyses used the medical procedure codes in combination with the diagnosis codes to obtain a more secure outcome definitions. The analysis of hip fractures relied on a combination of diagnosis and a hip fracture repair procedure to identify an outcome and the analysis of gastrointestinal bleeding used as an outcome definitions an endoscopy procedure in combination with a GI bleed diagnosis. Other outcome analyses used only diagnosis codes and the measured effects are explicitly defined in terms of rates of reported symptomologies.

### *Institutional Service Claims*

Medicaid and Medicare institutional claims are used in the outcome analyses as sources of diagnostic and procedure information. Institutional claims can be for inpatient stays or outpatient visits. The claims are generally not directly for medical procedures but for use of institutional facilities, for example inpatient accommodations or use of treatment rooms. The following key data elements are present on institutional claims: patient identifier, provider identifier, dates of service, diagnoses, procedure and/or revenue center code. Institutional claims for inpatient and outpatient services are used in the analyses in the same way as physician claims: for the formulation of covariate values and the determination of outcome status. Outpatient institutional claims are especially useful as a data source for care delivered by medical residents who are not allowed to bill directly for their services; in this context no physician claim is submitted but the institution does bill for facility use and supports the claim with diagnosis and procedure information. Inpatient claims distinguish primary diagnoses from other diagnoses reported in the course of the hospital stay. This distinction can be important when an outcome identifies an acute event as opposed to a simple report of disease presence. One outcome analysis, gastrointestinal bleeding, used the primary diagnosis as a flag for a severe outcome event.

## Analytic File and Variable Construction

The enrollment and claims databases are organized into a relational database keyed by individual patient identifiers. The outcome analyses were based on person summary records constructed from the main databases. The content of the person summary records varies for each outcome analysis in the covariate and outcome information captured, however the underlying design is the same across analyses. Baseline and follow-up periods were calculated using eligibility data from the enrollment file and demonstration contact data derived from drug claims. Outcome indicators were established through reviews of diagnostic and procedure data reported on claims with service dates in the follow-up period. Clinical covariate indicators were established through reviews of drug use, diagnostic and procedure data reported on claims with service dates in the baseline period. The person summary files were designed for easy analysis using standard computerized statistical processing tools.

## 4.8 The Impact of Project C.A.R.E. on Gastrointestinal Events

The HCFA DUR Demonstrations are designed to enhance Medicaid pharmacy practice through encouraging pharmacists to increase their review of patient drug therapies. The Washington State Demonstration, Project C.A.R.E., has focused on the pharmacist as a provider of medical care and has established a reimbursement mechanism for time spent in patient counseling — *cognitive services.* It was hypothesized that through providing reimbursement for cognitive services, the effectiveness of pharmacists as medical providers will be enhanced, more diagnostic information will be collected from clients and greater coordination between pharmacists and physicians will occur. Greater client contact will result in a more thorough identification of patients who are vulnerable to prescription-related adverse events.

Cognitive services provided by pharmacists can prevent a range of prescribing problems, including drug-drug interactions, drug-disease interactions, dosage errors, duplicative therapy, and patient noncompliance. The vulnerable populations served by Medicaid, which include indigent elderly people and children, are at higher risk for these prescribing problems and their associated adverse outcomes [Christensen et.al. (1996)]. These groups therefore are the most likely to benefit from an effective cognitive services program.

The overall impact of the programs was anticipated to be the more efficient and therapeutic utilization of prescription drugs and, potentially, a decrease in health outcomes linked to avoidable adverse effects from pharmaceutical exposures. The design of the C.A.R.E. program included the random assignment of volunteering pharmacies into two groups: 1) pharmacies that document cognitive services and receive payment; 2) pharmacies that document cognitive services but are not paid for those services. The division of participating pharmacies into payment and non-payment groups leads directly to an experimental design to test the impact of payment policy on health outcomes.

### Therapy-Related GI Ulcer Formation and Hemorrhage

Gastrointestinal symptoms are widely associated with exposures to pharmaceutical agents. The most severe outcomes are hospitalizations for upper GI ulcer formation and bleeding linked to the use of Non-Steroidal Anti-Inflammatory Drugs (NSAIDs). Non-aspirin NSAIDs are one of the most frequently prescribed drug categories and aspirin is a common over-the-counter medication. The high frequency of NSAID use, the severity of the health effects and the wide dissemination in the medical community of information about the association of NSAIDs to GI events, all suggest that pharmacy level interventions may be effective in reducing GI outcomes.

It is widely recognized that the risk of gastrointestinal ulceration and hemorrhaging (GI bleed) is significantly increased by exposure to aspirin and non-aspirin Non-Steroidal Anti-Inflammatory Drugs. Since the 1980s the medical literature has included numerous papers documenting the relationship between exposures of NSAIDs to hospitalization for GI bleeding, a selective search of Medline for articles on GI outcomes related to NSAID exposure yielded 44 major new studies and comprehensive reviews in the last 10 years in the major medical journals. A review of population cohort studies which evaluated the risk of GI events as stratified by NSAID exposure showed relative risks ranging from 1.4-2.1 [Carson et. al. (1994),3]. The risk of GI outcomes is elevated in populations which have histories of GI disease, other co-morbidities that relate to GI symptoms or pharmacy therapies that increase the risk of GI bleeding. Several treatment modalities have been developed to reduce the frequency of NSAID related GI disease in at risk populations, including the substitution of non-NSAID agents, reduction in NSAID dosage levels and prophylactic co-therapies. Pharmacist-based interventions that result in greater supervision of therapies may lead to reductions in the frequency of GI outcomes in a treated population.

Clinical and epidemiological studies of the relationship between NSAIDs and GI bleeding have used a variety of measures to determine the extent of the damage caused by the counter-prostaglandin action of NSAIDs. The studies have reviewed data ranging from survey data on GI symptoms to endoscopic evidence of tissue damage to mortality associated with severe episodes of GI bleeding. The population studies in the literature that use insurance claims data focus on hospitalization events and death, although several studies relied on outpatient diagnoses as well [Carson et. al. (1987),4]. Due to the relative rarity of hospitalization for ulcer and hemorrhage, there are four outcome types that are being studied, including outpatient encounters, with varying levels of specificity, severity and statistical power:

The analysis of the Demonstration's impact on the incidence of GI bleeding was based on the premise that improved pharmacy practices can be manifested as both short- and long-term effects through the modification of one or more prescription regimens and review of over-the-counter drug use. The underlying mechanism for reduction in outcome risk could be increased coordination between patients, pharmacists and prescribing physicians leading to more effective supervision of prescription drug regimens. This analysis focuses on the question of whether the use of Washington State C.A.R.E. pharmacies that are reimbursed for cognitive services was related to decreases in GI events and symptomology in client populations. The analysis measures the risk of a GI outcome in populations that are exposed to pharmacies that receive payments for cognitive services in comparison to populations exposed to pharmacies who receive no payments.

### Analytic Methods

The analyses measured the relationship between use of a demonstration pharmacy and the risk of a patient outcome. The models employed in these measurements rely on drug utilization and diagnostic information derived from Medicaid and Medicare pharmacy and medical claims. The conclusions of the analyses are adjusted for the demographic and disease covariates listed as relevant in the medical literature. In total the outcome analyses present three specific assessments of the efficacy of the C.A.R.E. program in improving the health of their patients according as measured by the odd ratio for GI outcomes in the payment and non-payment pharmacy groups.

- Physician or hospital diagnosis of an upper GI ulcer or hemorrhage;
- Physician or hospital diagnosis of an upper GI ulcer or hemorrhage and a GI endoscopy;
- Inpatient hospital diagnosis of upper GI ulcer or hemorrhage;
  - Inpatient primary hospital diagnosis of upper GI ulcer or hemorrhage;
  - Inpatient primary hospital diagnosis of upper GI ulcer or hemorrhage with subsequent death.

The objective of the analysis was to explore the relationship between payment for cognitive services and the rate of GI related adverse events. It was hypothesized that payment for cognitive services would increase the number of cognitive service interventions provided. This increase, in turn, was expected to change prescriber and/or patient behavior, and make prescriptions safer and more effective. Multiple logistic[28] regression was used in order to control for demographic, disease, and drug factors predisposing a given patient to the outcome under study. All control variables were measured during a baseline period before the demonstration began so as to be independent of any demonstration effects.

### *Study Population*

The analytic unit of observation was the Medicaid recipient who had contact with a demonstration pharmacy. Person level analyses are called for since the risk of outcome is primarily determined by individual disease and exposure covariates. Because the coverage of Medicaid managed care has been increasing over the period of the demonstration there will be portions of the population who will not be analyzable because of a lack of drug exposure and diagnostic data. Additionally Medicaid recipients who are institutionalized will not be directly subject to C.A.R.E. influences and will not be included in the models. As a result the analyses will focus on the non-nursing home Medicaid who remain under normal fee-for-service reimbursement coverage and who are users of C.A.R.E. pharmacies. Most Medicaid recipients who are permanently disabled or elderly will be Medicare entitled – dually eligible. Since Medicare is the primary payer of non-pharmacy services for these population, Medicare claims data is the source of health outcomes information and Medicaid is the sources of pharmacy use information.

### *Definitions of Baseline and Follow-up Periods*

Pharmacies enrolled in the demonstration in three waves (February, April, and September of 1994). Although other approaches would be possible, this study assumes that a patient's exposure to the demonstration (and consequently the patient's follow-up period) began when a patient first filled a prescription at a pharmacy participating in the demonstration. In the majority of cases, the follow-up period then continued until the demonstration ended in September 1995. In certain cases, if the patient lost eligibility for Medicaid, entered a nursing home, or entered an HMO, the follow-up period ended sooner because the outcomes data captured in the claims would be incomplete after that event. A patient's death was also deemed to terminate the follow-up period.

In a similar way, patients' baseline periods differed in length. The baseline period was determined to end the day before the patient's first contact with a demonstration pharmacy and to reach back twelve months or until the patient first became eligible for became eligible for the Medicaid program or left an HMO. The restrictions are imposed in order to ensure that the medical record is complete during the baseline period. Clinical outcomes that occurred on the day the patient first went to a demonstration pharmacy are excluded from both the baseline and the follow-up periods. In such cases, it is difficult to determine whether the clinical issue precipitated the prescription or the prescription led to the clinical issue without a medical records review.

---

[28] The SAS System's CATMOD procedure was used with the NOGLS option specified

### *Dependent Variables*

Based on the research literature, the GI events analyzed in previous studies relate to a hospitalization event with a upper GI ulcer or hemorrhage diagnosis [Carson et.al. (1987)]. Several studies narrowed the outcome definition and focused on death subsequent to a GI hospitalization [Griffin et. al. (1988); Guess et. al. (1988)]. Because of the increase in ambulatory surgical treatments, the outpatient or inpatient occurrence of a GI endoscopy with a GI ulcer or hemorrhage diagnosis was also identified as an analyzable outcome [Duggan et. al. (1986)]. The epidemiological evidence focused on measuring severe effects associated with NSAID exposure. The requirements of the analysis of the demonstration program argue for a broader review of the impact of a pharmacy intervention on both the incidence and prevalence of GI events and symptomology. The review of follow-up experience subsequent to the first use of a demonstration pharmacy differs in scope and nature from testing for severe GI events after an NSAID exposure. Improvements in pharmacy practice related to better supervision of NSAID use would be expected to be manifested in the reduction of severe GI outcomes, but also, more broadly, in the alleviation in GI distress as reflected in a reduction of reported GI symptoms as seen in outpatient diagnoses. The analyses reviewed outcomes that spanned the range of identifications of GI outcomes using the information available in claims data. The uncertainties in accurate diagnostic coding and the possible confounding relationship between demonstration pharmacy use and physician visit patterns for minor symptoms may serve to confound the most general measures.

**Exhibit 4.8.1**

**HCFA DUR Demonstration Evaluation**
**GI Outcome Markers**

| Outcome Type | Data Source | ICD9/CPT |
|---|---|---|
| Primary inpatient hospital diagnosis for ulcer or gastrointestinal hemorrhage followed by death within 30 days | Medicaid inpatient hospital primary diagnosis , Medicare inpatient hospital primary diagnoses and date of death from Medicaid eligibility records or Medicare enrollment records | ICD9 531-534, 578 |
| Primary inpatient hospital diagnosis for ulcer or gastrointestinal hemorrhage | Medicaid inpatient hospital primary diagnosis , Medicare inpatient hospital primary diagnoses | ICD9 531-534, 578 |
| Any inpatient hospital diagnosis for ulcer or gastrointestinal hemorrhage | Medicaid inpatient hospital diagnosis , Medicare inpatient hospital diagnoses | ICD9 531-534, 578 |
| Any hospital or physician diagnosis for ulcer or gastrointestinal hemorrhage with a GI endoscopy | Medicaid or Medicare hospital, outpatient or physician diagnosis and procedure codes | ICD9 531-534, 578<br>CPT 432 |
| Any hospital or outpatient diagnosis for ulcer or gastrointestinal hemorrhage | Medicaid or Medicare hospital, outpatient or physician diagnosis | ICD9 531-534, 578 |

*Sources:* Review of literature of GI events and NSAID exposure

Outcome indicators were constructed using Medicaid and Medicare diagnosis. Each outcome was associated with a series of ICD-9 codes. If a patient's record showed evidence of one of those ICD-9s as either the primary or secondary diagnosis for an inpatient stay or for a physician encounter (paid by Medicaid or Part A or Part B of Medicare) during the follow-up period, then the patient was determined to have had the outcome.

### *Independent Variables – Clinical Risk Factors*

Each model included a basic set of covariates that captured exposure to the intervention as well as general predisposing and need characteristics. Additionally, covariates were added to each model to control for clinical risk factors that are specific to the outcomes. The main set of independent variables, as well as the outcome specific independent variables, are described in the following sections.

Based on a review of the literature a broad selection of factors related to GI outcomes were identified. The major factors were related to a prior history of GI disease, disease that increase the risk of GI outcomes and conditions whose treatment is linked to increased risk of GI outcomes. Because the source of the covariate information was administrative claims data, diagnoses and pharmaceutical treatments related to the diagnoses were both used for the assessment of individual status.

Although the demonstration was designed as a random experiment, randomization occurred at the level of the pharmacy cluster while clinical benefits accrued at the level of the patient.[29] This posed several methodological challenges. First, individuals had the potential to go to multiple pharmacies and, potentially, to go to a mix of intervention, control, and non-demonstration pharmacies. The study adopted a simple approach. Individuals were considered members of the intervention group if they ever had a Medicaid-paid drug claim at an intervention pharmacy and never had such a claim at a control pharmacy. They were considered members of the control group if they ever had a Medicaid-paid drug claim at a control pharmacy and never had such a claim at an intervention pharmacy. Individuals who went to both intervention and control pharmacies were excluded from the study.[30]

The basic design of the study was to compare treatment and control patients within a regression framework that controls for demographic, drug, and disease factors. For all outcomes, regressions included controls for the patient's race, sex, and age. These factors affect health outcomes both directly and as proxies for behavior; they may also affect the predisposition to use care. These variables, which were taken from the enrollment records of the Medicaid program, were included in regressions as a series of indicator variables.

---

[29] Pharmacy clusters, built around common prescribers, were created and randomized to minimize the influence that intervention pharmacies might exert through common prescribers on control patients. In WA, there were 38 pharmacy clusters ranging in size from 1 person to 1200 people.

[30] For this purpose, the day that the patient first went to the demonstration pharmacy was included in the follow-up period.

**Exhibit 4.8.2**

### HCFA DUR Demonstration Evaluation Clinical Covariates[31]

| Group Name | Description |
|---|---|
| GI1 | physician or hospital; diagnosis for upper GI ulcer |
| | physician or hospital diagnosis for GI hemorrhage |
| | H2-antagonist and others |
| | antacids |
| | Misoprostol |
| GI2 | non-infective enteritis |
| | colitis |
| | diverticulitus |
| GI3 | disease of esophagus |
| | gastritis |
| | disorders of stomach |
| | gastroduodenal disorders |
| | other peritoneal disorders |
| CRD1 | hypertension |
| | anti-hypertensives |
| | beta-blockers |
| | calcium channel blockers |
| CRD2 | cardiotonics |
| | anti-anginals |
| | anti-arrhytmics |
| DB | diabetes |
| | insulin, oral hypoglycemics |
| CR | chronic renal failure |
| CP | chronic obstructive pulmonary disease/asthma |
| CS | use of corticosteroids |
| AN | anemia |
| CD | coagulation disorders |
| | anti-coagulant |
| AL | hepatic disease |
| | alcoholism |
| NS | NSAID |

*Sources:* Review of literature of GI events and NSAID exposure

### *Independent Variables – Length of Baseline and Follow-up Periods*

A patient with a longer follow-up period would be more likely to experience any clinical outcome than one with a shorter follow-up period, thus the length of the follow-up period was made an additional control in models. While the length of the baseline period does not affect patients' probabilities of clinical problems in and of itself, expected health spending, number of drugs, and risk factors will tend to be slightly under measured for patients with short baselines relative to patients with full baselines. Thus, controlling for measured covariates, the probability of adverse clinical outcomes might be higher for patients with short baselines; consequently, the length of the baseline period was included as an additional control. The lengths of the follow-up and baseline period were entered in models as the natural log of the number of months in each period.

---

[31] Diagnosis codes are specified to the third digit only Reviews of coding of the ICD9 manual and observed patterns suggest that broadly defined groups are appropriate for the conditions listed for covariates. The actual selection of drug covariates was based on NDC lists prepared by Pennsylvania State University research team.

60 and 75. The other rationale for an age limit is that the Medicaid program covers very specific populations, the younger groups are primarily women with young children who are not considered at risk for NSAID induced GI effects. The older populations, after the exclusion for nursing home residency, are a mixture of the community dwelling low income disabled and elderly.

**Exhibit 4.8.4**

**HCFA DUR Demonstration Evaluation**
**Study Populations**

| Study Population | Description |
|---|---|
| Main Population | Eligible for either Medicare and Medicaid, at least 45 years old in baseline period, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies. |
| NSAID User | Eligible for either Medicare and Medicaid, at least 45 years old in baseline period, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled an NSAID prescription in baseline period. |
| Any GI Diagnosis or Related Prescription | Eligible for either Medicare and Medicaid, at least 45 years old in baseline period, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, physician or hospital diagnosis of GI ulcer or hemorrhage or a prescription filled for a H2-Blocker, antacid or Misoprostol in the baseline period. |

*Sources:* Washington Project C.A.R.E. Medicare/Medicaid Files

## Results

The users of demonstration of pharmacies were assigned to payment and non-payment groups depending on the C.A.R.E. status of their pharmacy. The C.A.R.E. pharmacies were assigned to the payment (Group A) and non-payments groups (Group B) based on a cluster randomization methodology that was designed to minimize the overlaps in the physician-prescriber community potentially affected by demonstration influences. While the randomization was pharmacy based the outcomes analyses focused on individual patients. Below are tables of the population characteristics of the study groups in the baseline period for key covariates and the results of demographic and disease adjusted analyses.

**Exhibit 4.8.5**

**HCFA DUR Demonstration Evaluation**
**Demonstration Descriptive Statistics**

| Selected Baseline Statistics | Demonstration Group A | Demonstration Group B | T-Test Probability Prob>ITI Unequal |
|---|---|---|---|
| Study Subjects | 17,639 | 11,932 | |
| Person-Months | 260,515 | 176,346 | |
| Mean Age | 62.2 | 63.0 | 0.0001 |
| Mean Gender (0=M/1=F) | 0.65 | 0.67 | 0.0002 |
| Mean Baseline Months | 10.0 | 9.8 | 0.0001 |
| Mean Follow-up Months | 14.8 | 14.4 | 0.0001 |
| Primary Diagnosis of Gastric Ulcer | 0.3% | 0.3% | 0.5641 |
| Primary Diagnosis of GI Hemorrhage | 0.2% | 0.3% | 0.0699 |
| Diagnosis of Gastric Ulcer | 3.3% | 3.8% | 0.0147 |
| Diagnosis of GI Hemorrhage | 3.1% | 3.5% | 0.0361 |
| GI endoscopy Procedure | 2.3% | 2.4% | 0.4401 |
| Arthritis Diagnosis | 38.0% | 37.4% | 0.3367 |
| Diabetes Diagnosis | 15.2% | 16.1% | 0.0323 |
| COPD/Asthma Diagnosis | 19.5% | 20.0% | 0.2810 |
| CRF Diagnosis | 1.8% | 1.6% | 0.1973 |
| Alcoholism Diagnosis | 4.0% | 3.2% | 0.0001 |
| Hypertension Diagnosis | 28.1% | 26.3% | 0.0006 |
| NSAID Rx | 38.1% | 34.0% | 0.0001 |
| H2-Blocker Rx | 29.0% | 28.7% | 0.5992 |
| Misoprostol Rx | 1.6% | 1.3% | 0.0140 |
| Antacid Rx | 7.5% | 5.3% | 0.0001 |
| Corticosteroid Rx | 10.6% | 10.7% | 0.7585 |
| Anti-Coagulant Rx | 3.5% | 3.9% | 0.1387 |

*Sources:* Washington Project C.A.R.E. Medicare/Medicaid Files

### *Baseline Statistics*

The use of NSAIDs in the baseline period coupled with a GI diagnosis or another drug therapy that can increase the risk of GI bleeding may be indicative of pharmacy practices that can be improved. The statistics in Exhibit 4.8.6 are measures of problematic sub-populations as distributed by their affiliation with Project C.A.R.E. pharmacy types at the time of demonstration induction. Similarly, the rate of PSU/UMPCPS screen violations for NSAID therapies can be used to identify high dosage or duplicate

**Exhibit 4.8.6**

**HCFA DUR Demonstration Evaluation**
**Diagnosis**

| NSAID-GI Baseline Statistics | Demonstration Group A (N=17,639) | Demonstration Group B (N=11,932) | T-Test Probability Prob>ITI Unequal |
|---|---|---|---|
| NSAID - Ulcer or Hemorrhage | 1.5% | 1.5% | 0.8410 |
| NSAID - H2-Blocker Rx | 11.5% | 10.9% | 0.1219 |
| NSAID - Antacid Rx | 3.0% | 2.0% | 0.0001 |
| NSAID - Misoprostol Rx | 1.2% | 0.9% | 0.0141 |
| NSAID - Anti-Coagulant Rx | 0.7% | 0.8% | 0.5528 |
| NSAID - Corticosteroid Rx | 3.0% | 3.1% | 0.8170 |
| NSAID PSU Dosage Violation | 14.3% | 13.6% | 0.0989 |

*Sources:* Washington Project C.A.R.E. Medicare/Medicaid Files

regimes. The statistics are measures of concurrent prescriptions filled in the same month and diagnoses reported in the same month as a NSAID prescription fill or in the subsequent month.

The populations assigned to the two pharmacy types are generally similar in personal demographics and in the length of baseline and follow-up periods, although the size of the study groups leads to statistical significance for small differences. There are significant differences in prevalence of NSAID use and the frequency of GI bleed diagnoses. The overall rate of NSAID prescribing is lower during the baseline period in Group B and the frequency of GI diagnoses is higher — although the prevalence of H2-blocker utilization is the same. The portion of the population with an endoscopy procedure in the baseline period is similar in the two groups. There are elevated rates of antacid and Misoprostol therapies concurrent with NSAID prescriptions in the treatment pharmacies.

In summary the treatment patients appear to have elevated rates of antacid, Misoprostol and NSAID therapies and a decreased rates of GI bleed diagnoses in all settings. In contrast hospitalization with a primary diagnosis of ulcer or hemorrhage are equally distributed, as are endoscopy procedures, GI bleed diagnosis concurrent to an NSAID prescription and concurrence between NSAID therapy and H2-blocker use. This may be suggestive of greater GI prophylaxis with higher NSAID utilization in the treatment pharmacies but greater rates of reported GI bleed diagnoses in the non-treatment patients. Indications of severe GI effects, i.e. hospitalizations and endoscopy procedures are not significantly different in the two populations.

### *Regression Results*

In total five models were prepared for the main population and two risk groups, resulting in ten regression analyses. The parameter estimates for all variables used in the logistic regression of the main population are included in Appendix B. The demographic and clinical covariates generally showed the anticipated relationship to GI outcomes. Three factors that consistently showed a non-significant relationship to any GI outcomes were withdrawn from the models, they were: 1) exposure to an anti-coagulant medication or an coagulation disorder diagnosis in the baseline period; 2) the drugs grouped under CRD2; 3) the diagnoses grouped under GI3.

Below are presented the adjusted odds ratios derived from the covariance matrix for each model in the general population and two of the risk groups. The main covariate influences were observation time, longer follow-up leads to greater risk, increased age is correlated with greater risk of GI outcomes, gender, female reduces risk and a baseline GI diagnosis is highly associated with the occurrence of events in the study period. Baseline diagnoses of renal failure, anemia, alcoholism were all strongly correlated with GI bleed outcomes.

**Table 4.8.7**

**HCFA DUR Demonstration Evaluation**
**Estimated Odds Ratios: Major Dependent Variables, Adjusted for Follow-up Differences in Characteristics**

| General Population Age>45 | OR | 95% CI |
|---|---|---|
| Hospitalization with a primary diagnosis of ulcer or GI | 3.52 | 0.51-24.34 |
| Hospitalization with a primary diagnosis of ulcer or GI | 0.97 | 0.76-1.25 |
| Inpatient diagnosis of ulcer or GI hemorrhage | 0.90 | 0.76-1.06 |
| Inpatient or outpatient diagnosis of ulcer or GI hemorrhage and | 1.04 | 0.88-1.22 |
| Inpatient or outpatient diagnosis of ulcer or GI hemorrhage | 0.88 | 0.81-0.96 |
| **Population Age>45, NSAID User in Baseline Period,** | **OR** | **95% CI** |
| Hospitalization with a primary diagnosis of ulcer or GI | – | – |
| Hospitalization with a primary diagnosis of ulcer or GI | 0.97 | 0.65-1.46 |
| Inpatient diagnosis of ulcer or GI hemorrhage | 0.89 | 0.68-1.16 |
| Inpatient or outpatient diagnosis of ulcer or GI hemorrhage and | 0.90 | 0.69-1.16 |
| Inpatient or outpatient diagnosis of ulcer or GI hemorrhage | 0.79 | 0.70-0.90 |
| **Population Age>45, Any GI Ulcer/Hemorrhage Diagnosis, H2-** | **OR** | **95% CI** |
| Hospitalization with a primary diagnosis of ulcer or GI | 0.37 | 0.01-12.37 |
| Hospitalization with a primary diagnosis of ulcer or GI | 0.93 | 0.67-1.31 |
| Inpatient diagnosis of ulcer or GI hemorrhage | 0.86 | 0.69-1.07 |
| Inpatient or outpatient diagnosis of ulcer or GI hemorrhage and | 0.86 | 0.70-1.06 |
| Inpatient or outpatient diagnosis of ulcer or GI hemorrhage | 0.83 | 0.75-0.93 |

*Sources:* Washington Project C.A.R.E. Medicare/Medicaid Files

The adjusted odds ratios derived from the regression parameter estimates show no significant result for Project C.A.R.E. pharmacies in the A group in decreasing the risk of severe GI outcomes. The statistical power for determining the effect of payment for cognitive services on death subsequent to a hospitalization for GI ulcer or hemorrhage is insufficient in the study population for any definitive measurement of risk. Hospitalizations with a primary diagnosis of GI bleed were not significantly decreased. Reports of GI bleed diagnoses during inpatient stays had non-significant decreases in risk in all models. Endoscopy procedures were not significantly reduced.

The only consistent positive result was an apparent significant decrease in the reporting of a GI bleed diagnosis in all settings in the general population and in two of the risk groups (NSAID users in the baseline and the population with a prior indication of GI symptoms). The mechanism for a decrease in risk for general GI symptomology could follow two tracks. The first is that changes is therapeutic prescribing regimens resulted in lower rates of adverse effects. The second track is that the symptomology remained the same, however pharmacy base cognitive services reduced the need for physician consults. In either case the analysis suggests a possible benefit in patient surveillance for GI problems related to C.A.R.E. pharmacy status.

### Discussion

There are external and internal factors that could explain the decrease in GI diagnoses. An important explanatory element would be the connection between modifications in patient drug therapies related to the decrease in risk. Because many of the important drugs are over-the-counter some of the most key modifications, e.g. elimination of aspirin or non-prescribed antacids, are essentially unobservable. Additionally diagnostic information is sensitive to physician reporting practices and since physicians were not randomized there may be biases in the distribution of physician specialties and practice patterns in the experimental groups that would cause differential reporting of GI diagnoses.

Below in Exhibit 4.8.8 we present an analysis of several therapy modalities which are considered to be problematic — NSAID with concurrent corticosteroid, NSAID with an anti-coagulant, PSU/UMPCPS NSAID screener violation — as well as statistics relating to the populations using NSAIDs, H2-blockers, antacids and Misoprostol in the follow-up period. The measurements are based on binary variables of the presence or absence of drug use in the follow-up period. The statistics are measures of concurrent prescription dispense events in the same month and diagnoses in the same month or in the subsequent month.

**Exhibit 4.8.8**

**HCFA DUR Demonstration Evaluation**
**NSAID Drug Utilization: Project C.A.R.E. Treatments (Group A) and Controls (Group B)**

| NSAID-GI Follow-up Statistics | Demonstration Group A (N=17,639) | Demonstration Group B (N=11,932) | T-Test Probability Prob>ITI Unequal |
|---|---|---|---|
| NSAID - Ulcer or Hemorrhage | 2.6% | 2.5% | 0.8302 |
| NSAID - H2-Blocker Rx | 14.8% | 13.9% | 0.0369 |
| NSAID - Antacid Rx | 3.5% | 2.3% | 0.0001 |
| NSAID - Misoprostol Rx | 1.4% | 0.8% | 0.0001 |
| NSAID - Anti-Coagulant Rx | 1.3% | 1.1% | 0.0735 |
| NSAID - Corticosteroid Rx | 4.7% | 4.6% | 0.6426 |
| NSAID PSU Dosage Violation | 14.0% | 12.8% | 0.0021 |

*Sources:* Washington Project C.A.R.E. Medicare/Medicaid Files

A logistic regression analysis using the same covariates as for the analysis of GI effects shows significant elevations in Misoprostol and antacid therapy concurrent with an NSAID prescription in the treatment pharmacies during the follow-up period. Concurrent NSAID and H2-blocker therapies were not significantly different in the two groups after age, gender and clinical covariates were adjusted for. The elevated use of antacids and Misoprostol among treatment patients is significant, however the differences are of the same magnitude as observed in the baseline data, corrected for the longer follow-up observation period.

The randomization of patients based on minimization of overlap in prescriber communities could have resulted in biases in the reporting and treatment of GI symptoms. The use of diagnoses from all settings as a method for measuring disease prevalence can be problematic if there are strong community variations in the coding and reporting of secondary symptoms that result in observation bias. Outcomes which are not well defined events are especially sensitive to these factors.

Similarly, community practices of prescribers and pharmacies can vary during the baseline and evolve differentially over the follow-up period, independent of the effect of the C.A.R.E. program. Changes in community standards that do not equally affect the study populations cannot easily be adjusted for, however pre-existing differences can be measured in a review of the baseline data.

An analysis corrected for demographic characteristics of differences in the two study groups during the baseline and follow-up periods is presented below. The statistic is listed as the odds ratio for GI events prior to the demonstration by study group adjusted by demographic and observation time factors only. The question that is addressed is whether the observed effect in the treatment population pre-exists the start of the demonstration.

**Exhibit 4.8.9**

**HCFA DUR Demonstration Evaluation**
**Estimated Odds Ratios: Major Dependent Variables, Adjusted for Baseline and Followup Differences in Characteristics**

| General Population Age>45 | OR | 95% CI |
|---|---|---|
| Hospitalization with a primary diagnosis of ulcer | 0.94 | 0.69-1.29 |
| Inpatient diagnosis of ulcer or GI hemorrhage | 0.80 | 0.65-0.98 |
| Inpatient or outpatient diagnosis of ulcer or GI | 0.96 | 0.77-1.19 |
| Inpatient or outpatient diagnosis of ulcer or GI | 0.84 | 0.77-0.93 |
| **General Population Age>45** | **OR** | **95% CI** |
| Hospitalization with a primary diagnosis of ulcer | 1.01 | 0.78-1.30 |
| Inpatient diagnosis of ulcer or GI hemorrhage | 0.92 | 0.78-1.08 |
| Inpatient or outpatient diagnosis of ulcer or GI | 1.06 | 0.90-1.25 |
| Inpatient or outpatient diagnosis of ulcer or GI | 0.89 | 0.83-0.97 |

*Sources:* Washington Project C.A.R.E. Medicare/Medicaid Files

The odds ratio for the treatment group for GI bleed diagnosis – adjusted for age, gender, race and observation time – in the baseline and follow-up periods is essentially unaltered. The following circumstances contribute to a concern that the measurement of a protective effect on GI symptomology for C.A.R.E. payment pharmacies is artifactual and linked to randomization methodology: 1) the lack of a clear measurement of changes in sensitive combined therapies, 2) no substantial decrease in the presence of PSU/UMPCPS NSAID screen violations, and 3) the observation of higher reports of GI ulcer and hemorrhage diagnoses in the baseline period. Lacking any significant measurement no substantive result can be attributed to cognitive services payment in the Washington demonstration on GI events of symptomology.

## 4.9 The Impact of the Cognitive Services Demonstration on Adverse Clinical Outcomes Related to the Use of Antidepressants

Payment for cognitive services is intended to encourage pharmacists to spend more time ensuring prescription quality and educating patients regarding the best way to take their drugs and avoid potential side effects. These efforts, in turn, are expected to lead to improvements in health outcomes.

This study tested the impact of Project C.A.R.E., a demonstration of payment for pharmacists' cognitive services that took place in Washington State in 1994 and 1995, on certain adverse clinical outcomes known to be related to less-than-optimal prescriptions and drug utilization. Specifically, it explored impacts on cardiac conditions, orthostatic hypotension (sometimes marked by hip fractures), and a cluster of anticholinergic symptoms; the clinical literature indicates that all of these conditions are associated with the use of antidepressant drugs. The main population of interest in this study consisted of elderly individuals who were eligible for both Medicaid and Medicare, and who went to pharmacies that were participating in the demonstration.

The results of the analysis suggested that Project C.A.R.E. did *not* have a measurable impact on adverse outcomes related to antidepressant drug use. Holding baseline characteristics constant, individuals who visited pharamacies where cognitive service events were reimbursed were just as likely as control group patients to experience cardiac effects, hypotension and anticholinergic symptoms. These findings suggest that while Project C.A.R.E. may have led to benefits in other areas, the intervention was not effective in improving antidepressant drug prescribing quality to the extent that major, adverse health outcomes were prevented.

### Adverse Clinical Outcomes Related to Antidepressants

Numerous studies support the conclusion that the use of tricyclic antidepressants (TCAs) at high doses is associated with serious cardiac effects. This class of drugs can affect the conduction system, increase the heart rate, and exert an effect on heart rhythm. Some researchers have attempted to identify whether or not these effects are observed at therapeutic doses as well as high doses. The research literature also links TCAs with orthostatic hypotension and anticholinergic effects.

All the outcomes under study are associated with tricyclic antidepressants, and not with other classes of antidepressants such as MAO inhibitors or targeted antidepressants, such as SSRIs. Initially, the outcomes were identified based on associations with the PSU screening criteria; these criteria emphasize TCAs. Also, because TCAs were less costly than SSRIs during the study period, they probably were the dominant class of antidepressant among the Medicaid population

*Cardiac Effects:*

***Conduction Disorders.*** TCAs can prolong cardiac conduction, but this event will almost never produce symptomatic heart disease at therapeutic levels for patients without pre-existing conduction disorders [Jackson et al. (1987)]. By slowing intraventricular conduction, TCAs can lead to complete heart block or ventricular reentry arrhythmias in cases of overdose or among patients with pre-existing cardiac disease [Glassman and Preud'homme (1993)]. At overdose levels, conduction delays and heart block can be significant, even fatal problems. For patients who already have conduction disorders, the conduction delays produced by TCA use can lead to additional cardiac problems, including fatal heart block [Roose et al.(1987)]. Among patients taking TCAs, those with pre-existing bundle branch blocks are thought to be at greatest risk of developing high degree atrioventricular conduction delays [Jackson et al. (1987)]. Jackson et al. (1987) state: "Apart from orthostatic hypotension, the major risk of the tricyclic antidepressants is tricyclic-induced atrioventricular block in patients with pre-existing conduction disease."

***Heart Rhythm Effects.*** TCAS have a class IA antiarrhythmic action, and several studies have indicated that drugs with class IA and/or IC antiarrhythmic action increase the risk of mortality among patients with arrhythmias [(Glassman (1993); Morganroth et al. (1991)]. Glassman (1993) further states that there has been recent evidence from additional studies that antiarrhythmic drugs may pose a mortality risk when used to suppress atrial fibrillation, as well as when used to suppress ventricular arrhythmias after MI.

The mechanism by which class I antiarrhythmic drugs produce increased mortality is not known [Glassman (1993)]. However, it appears that a likely candidate is an interaction between a class IA antiarrhythmic drug and an ischemic myocardium. Glassman and Roose (1994) present the current understanding of this relationship, "Class I antiarrhythmic drugs are strongly antiarrhythmic under aerobic conditions. However, during periods of oxygen deprivation, such as angina or myocardial infarction, these drugs appear to become paradoxically proarrhythmic."

The proarrhythmic effect can be expected to be seen in approximately 10 percent of patients with pre-existing arrhythmias who are given class I antiarrhythmics [Glassman and Roose (1994)]. Glassman and Roose go on to warn, however, that it is not only patients with pre-existing arrhythmia who have ischemic hearts. Avoiding TCA use in patients with arrhythmia will not protect patients with asymptomatic arrhythmias or ischemic heart disease but no arrhythmia. They further say that from a point of view of safety, it would be best to avoid a TCA "in any patient who is at serious risk for ischemic heart disease."

### Orthostatic Hypotension

Orthostatic hypotension is thought to be the most common of the serious side effects of TCA use at therapeutic levels. It is almost universally reported in articles written about the side effects associated with TCAs. Roose et al. (1987) state that orthostatic hypotension occurs in 5 percent to 10 percent of patients taking TCAs. In depressed patients who have impaired left ventricular function and are treated with imipramine the rate of orthostatic hypotension can be much higher [Glassman et al. (1983); Chutka (1990)].

Consequences of hypotension include falls and fractures; as Glassman (1984) points out, major drops in blood pressure occur at all ages, but the risk of serious sequelae increases considerably with age. Two studies by Ray et al. (1987 and 1991) examine the association of hip fracture (one of the major serious sequelae in the elderly) and TCA use. Ray et al. (1987) found that the relative risk of hip fracture for TCAs was 1.9 among the Medicaid population (95 percent C. I. 1.3-2.8). Ray et al. (1991) found the relative risk for current TCA use was 1.6 among an elderly Canadian population (95 percent CI 1.3-1.9). This study indicated that the risk of hip fracture was greatest in the first 90 days of therapy, though there was increased risk for longer durations of use as well.

### Anticholinergic Effects

TCAs also produce anticholinergic effects, the most common of which are dry mouth, blurred vision, constipation, tachycardia, and urinary hesitancy/retention [Coccaro and Siever (1985)]. A number of studies and reviews discuss the occurrence of anticholinergic effects. Perhaps the most recent is an article by Simon et al. (1996) in which fluoxetine (an SSRI) is compared with imipramine (a tertiary amine TCA) and desipramine (a secondary amine TCA) using clinical, functional, and economic outcomes. Constipation, GI upset, headache, stimulation and dizziness occurred more frequently among TCA patients.

### Evaluation Approach

This study investigated whether or not payment for cognitive services led to decreased rates of outcomes known to be associated with the use of antidepressants. For such an effect to exist, several things must have happened. First, the existence of per-service payments must have led pharmacists to perform more cognitive services; even without payment, pharmacists make efforts to guarantee that Medicaid recipients receive safe and effective drug therapy. Second, those additional services must either have changed prescriber behavior, possibly changing the nature of prescriptions written, or changed patient behavior, possibly improving compliance or awareness of adverse reactions, in ways that prevented adverse outcomes. Note that a pharmacist's contact with a physician regarding one patient may affect that physician's awareness and behavior in ways that benefit additional patients as well.

The basic design was to compare rates of clinical outcomes among patients who went to intervention pharmacies to those among patients who went to control pharmacies. Multiple regression was used in order to control for demographic, disease, and drug factors predisposing a given patient to the outcome under study. All control variables were measured during a baseline period before the demonstration began so as to be independent of any demonstration effects. A version of White's method that incorporates cluster based sampling was used to estimate standard errors [Rogers (1993)].

### Data

The data used in this study was extracted from Medicaid drug claims and Medicare health services utilization claims for patients in the state of Washington in 1993 (to provide a baseline) and 1994-1995. (The demonstration took place between February of 1994 and September of 1995.) Other records and data streams generated by the demonstration were also incorporated into the data set. Data development is described in detail in Section 4.7.

#### *Study Populations*

As displayed below in Exhibit 4.9.1, the data was used to create three study populations: (1) a main study population; (2) a sub-sample of the main study population consisting of those who had been using anti-depressants during the baseline period; and (3) a sub-sample of the main population consisting of those who had been using anti-depressants during the baseline or the follow-up period. Although all models were implemented using these three samples, not all results are presented in this chapter. Instead, this chapter focuses on interpreting and summarizing the results from these models; selected models estimated using the main population are presented in full in the appendix to this chapter.

**Exhibit 4.9.1**

**HCFA DUR Demonstration Evaluation**
**Antidepressant Study Populations**

| Study Population: | Description: |
|---|---|
| Main Population<br>(N = 7805) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies |
| Baseline or Follow-up Antidepressant Users<br>(N = 2121) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled at least one prescription for an anti-depressant during the baseline period or during the follow-up period |
| Baseline Antidepressants Users<br>(N = 1391) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled at least one prescription for an anti-depressant during the baseline period |

*Sources:* Washington Project C.A.R.E. Medicare/Medicaid Files

The main study population consisted of elderly patients (65 and over on January 1, 1993) who were eligible for both the Medicaid and Medicare programs and who went to either intervention or control pharmacies during the intervention. The elderly are particularly at risk for both the cardiac and hip fracture outcomes. Moreover, they are generally in more delicate health and taking more drugs than younger patients; the elderly thus are hypothesized to be more likely than people of other age groups to benefit from the intervention. The requirement of dual eligibility ensures that the analyses have complete and consistent records of drug and health services utilization for all patients in the data set.

The study population was further limited to individuals who used antidepressants during the baseline period, and individuals who used antidepressants during the baseline or follow-up periods. Narrowing the population in this way was intended to improve the specificity of the analysis.[32] The population of all antidepressant users (rather than TCA users alone) was used to retain patients for whom the intervention effect involved the substitution of another antidepressant for a TCA in order to avoid the clinical risks associated with TCAs. Furthermore, the analysis considered both a sub-sample of baseline anitdepressant users as well as a sub-sample of baseline or follow-up antidepressant users because of concerns that the intervention might affect anti-depressant use in the follow-up period. In this case, estimates based on the baseline/follow-up anti-depressant sub-sample might be biased.[33]

### *Definitions of Baseline and Follow-up Periods*

Pharmacies enrolled in the demonstration in three waves (February, April, and September of 1994). Although other approaches would be possible, this study assumes that a patient's exposure to the demonstration (and consequently the patient's follow-up period) began when a patient first filled a prescription at a pharmacy participating in the demonstration. In the majority of cases, the follow-up period then continued until the demonstration ended in September 1995.[34] In certain cases, if the patient lost eligibility for Medicare or Medicaid, entered a nursing home, or entered an HMO, the follow-up period ended sooner because the outcomes data captured in the claims would be incomplete after that event.[35] A patient's death was also deemed to terminate the follow-up period.

In a similar way, patients' baseline periods differed in length. The baseline period was determined to end the day before the patient's first contact with a demonstration pharmacy and to reach back twelve months or until the patient first became eligible for the Medicare program, became eligible for the Medicaid program, left an HMO, or left a nursing home and became resident in the community. Again, the latter three restrictions are imposed in order to ensure that the medical record is complete during the baseline period.

---

[32] Nevertheless, analyses based on this more narrow population would miss any intervention effect that involved patients avoiding the use of antidepressants altogether.

[33] If Project C.A.R.E. affected antidepressant use in the follow-up period and the population of interest consists of antidepressant users in the baseline and follow-up periods, a selection problem would exist. In this case, it's possible that those who benefitted from the intervention would be more likely than other to be excluded from the sample.

[34] In principle, effects of the intervention might last beyond the demonstration period. However, limited payment for CS was made available to both intervention and control pharmacies after September of 1995 which could confound any estimated intervention effects based on outcomes after that date.

[35] Nursing home prescriptions that were batched and sent to demonstration pharmacies logically could not be affected by cognitive services interventions.

Note that clinical outcomes that occurred on the day the patient first went to a demonstration pharmacy are excluded from both the baseline and the follow-up periods. In such cases, it is impossible to know whether the clinical issue precipitated the prescription or the prescription led to the clinical issue. The safe decision therefore was to omit the outcome from both periods.

### *Dependent Variables*

Based on the research literature, the most relevant aspects of which were summarized above, the following five clinical outcomes were selected for this study: (1) conduction disorders; (2) cardiac dysrhythmia; (3) hypotension[36]; (4) hip fractures; and (5) anticholinergic effects. These dependent variables are summarized below in Exhibit 4.9.2.

**Exhibit 4.9.2**

**HCFA DUR Demonstration Evaluation**
**Dependent Variables**

| Variable: | ICD-9 Classification: |
|---|---|
| conduction disorders | 426.0 - 426.9 |
| cardiac dysrhythmia | 427.0 - 427.9 |
| hypotension | 458.0 - 458.9 |
| new hip fractures | 820.0 - 820.9, also used CPT code |
| anticholinergic effects (any one of blurred vision, constipation, drowsiness, dry mouth, or urinary retention) | 368.8, 564.0, 780.09, 527.7, 788.2 |

*Sources:* International Classification of Diseases, 9th Revision, DHHS Publication No. (PHS) 91-1260, October, 1991.

Conduction disorders and cardiac dysrhythmia are clear choices, although the ambiguity regarding mechanism of action is troubling. Orthostatic hypotension was also a natural choice, and hip fractures offer a clinically significant marker for that condition. Finally, the study analyzed a cluster of five anticholinergic effects (blurred vision, constipation, drowsiness, dry mouth, and urinary retention). This analysis is the most problematic because these effects may not be clinically significant enough to guarantee that physicians note them in patients' records or even that patients report them to physicians.

---

[36] Coding practice often does not distinguish between hypotension and orthostatic hypotension so a broad outcome encompassing either was created.

Outcome indicators were constructed using Medicare diagnosis codes.[37] Each outcome was associated with a series of ICD-9 codes. If a patient's record showed evidence of one of those ICD-9s as either the primary or secondary diagnosis for an inpatient stay or for a physician encounter (paid by Part A or Part B of Medicare) during the follow-up period, then the patient was determined to have had the outcome.

For each outcome, a model also was estimated in which the outcome variable was based on inpatient claims alone. It is possible that compared to physicians in the outpatient setting, hospitals might be more precise in their coding of diagnoses because of its payment implications. Variables that were created using inpatient claims only therefore might contain less measurement error compared to variables created using all claims.

In the new hip fracture model, the use of variables created using inpatient claims only was particularly useful because physician claims might capture follow-up care rather than the hip fracture event itself. Hip fracture CPT codes for the main population were compared with inpatient claims for hip fracture to assure that the inpatient claims were capturing the hip fracture event. The rate of hip fracture based on CPT codes and the rate of hip fracture based on ICD-9 codes and inpatient claims were similar. This similarity suggests that the hip fracture variable based on ICD-9 codes and inpatient claims is a good measure of the new hip fracture event.[38]

### *Independent Variables*

Each model included a basic set of covariates that captured exposure to the intervention as well as general predisposing and need characteristics. Additionally, covariates were added to each model to control for clinical risk factors that are specific to each of the outcomes. The main set of independent variables, as well as the outcome specific independent variables, are described in the following sections.

***Exposure to the Intervention***. Although this demonstration was designed as a random experiment, randomization occurred at the level of the pharmacy cluster while clinical benefits accrued at the level of the patient.[39] This posed several methodological challenges. First, individuals had the potential to go to multiple pharmacies and, potentially, to go to a mix of intervention, control, and non-demonstration pharmacies. The study adopted a simple approach. Individuals were considered members of the intervention group if they ever had a Medicaid-paid drug claim at an intervention pharmacy and never had such a claim at a control pharmacy. They were considered members of the control group if they ever had a Medicaid-paid drug claim at a control pharmacy and never had such a claim at an intervention pharmacy. Individuals who went to both intervention and control pharmacies were excluded from the study.[40]

---

[37] Because all of the samples used in this analysis included individuals who were dually eligible for Medicare and Medicaid, it is assumed that all medical services were initially billed and primarily covered by Medicare.

[38] In the main population in WA, there were 103 hip fractures recorded using CPT code information, 130 hip fractures recorded using inpatient Medicare claims, and 187 hip fractures recorded using all Medicare claims. The higher rate of hip fractures recorded using all claims probably exists because outpatient, follow-up care for an old hip fracture has been coded as hip fracture.

[39] Pharmacy clusters, built around common prescribers, were created and randomized to minimize the influence that intervention pharmacies might exert through common prescribers on control patients. In WA, there were 38 pharmacy clusters ranging in size from 1 person to 1200 people.

*General Predisposing and Need Characteristics.* The basic design of the study was to compare treatment and control patients within a regression framework that controls for demographic, drug, and disease factors. For all five outcomes, regressions included controls for the patient's race, sex, and age. These factors affect health outcomes both directly and as proxies for behavior; they may also affect the predisposition to use care. These variables, which were taken from the enrollment records of the Medicaid program, were included in regressions as a series of indicator variables.

### Length of Baseline and Follow-up Periods

Other things equal, a patient with a longer follow-up period would be more likely to experience any clinical outcome than one with a shorter follow-up period, thus the length of the follow-up period was made an additional control in models. While the length of the baseline period does not affect patients' probabilities of clinical problems in and of itself, expected health spending, number of drugs, and risk factors will tend to be slightly under measured for patients with short baselines relative to patients with full baselines. Thus, controlling for measured covariates, the probability of clinical outcomes might be higher for patients with short baselines; consequently, the length of the baseline period was included as an additional control. The lengths of the follow-up and baseline period were entered in models as the natural log of the number of months in each period.

### Pharmacy Wave

Pharmacies entered the cognitive services demonstration in three waves. The first wave of 160 pharmacies enrolled in February 1994. A second wave (N = 33) and a third wave (N = 7) of pharmacies entered in April and September 1994. It is possible that pharmacies entering the demonstration later learned from the experiences of prior entrants and provided more effective cognitive services. Additionally, the rate of cognitive service events per 100 dispensed prescriptions varied widely over time in the treatment group. For these reasons, indicator variables were included in the models that capture the wave of the first demonstration pharmacy where the individual filled a prescription. The omitted category in the models was wave one.

The basic set of independent variables that were used in every model (as described above) are summarized below in Exhibit 4.9.3.

---

[40] For this purpose, the day that the patient first went to the demonstration pharmacy was included in the follow-up period.

**Exhibit 4.9.3**

**HCFA DUR Demonstration Evaluation**
**Basic Set of Covariates Included in All Models**

race
sex
age (in categories)
length of baseline period
length of follow-up period
wave of first contact demonstration pharmacy

### Risk Factors Specific to Each Outcome

*Conduction Disorders/Cardiac Dysrhythmia*. Analyses of these outcomes added additional controls for pre-existing conduction disorders, cardiac dysrhythmia, heart failure, and other cardiac diagnoses. These conditions are risk factors for ongoing conduction disorders and cardiac dysrhythmia. The models also controlled for the use of antiarrhythmics, beta blockers, CCBs, and digoxin in the baseline period because these drugs also might indicate pre-existing arrhythmia. Finally, analyses controlled for the use of TCAs during the baseline period; note that even if the main effect of the intervention were to take patients off TCAs, the change would be observed in the follow-up, not the baseline period.

*Hypotension.* Analyses of this outcome controlled for pre-existing hypotension, the use of antihypertensives (which sometimes induce hypotension), TCAs, and low potency antipsychotics (also a risk factor) during the baseline period.

*Hip Fracture.* Analyses of this outcome controlled for pre-existing hip fracture, hypotension, gait abnormality and osteoporosis. These conditions are known to raise the risk of hip fractures. In addition, the models controlled for the use of a number of drugs during the baseline: (1) sedative hypnotics or benzodiazepines, which may make patients drowsy and raise the risk of falls; (2) antihypertensives; (3) corticosteroids (chronic use) which potentially induce hypotension; and (4) TCAs.

*Anticholinergic Effects.* Analyses of this outcome controlled for pre-existing anti-cholinergic effects, pre-existing diabetes (which may cause anticholinergic effects) and for incontinence and constipation (which show evidence of ongoing problems). The models also controlled for baseline use of a number of drugs known to induce anticholinergic effects: (1) antihistamines; (2) antispasmodics; (3) benzodiazepines; (4) narcotic analgesics; (5) low potency antipsychotics; and (6) TCAs.

### Rate of Clinical Outcome in Pharmacy Cluster

The pharmacy cluster, the unit of randomization in this analysis, effectively defined medical communities. Potentially, there could be unmeasured differences in demographic and behavioral characteristics among patients in different pharmacy clusters that influenced the need for care and the predisposition to seek it. Similarly, there could be geographic variation associated with cluster in access to care and in its practice, such as variations in coding of conditions and diagnoses and in treatment protocols. If these differences were associated with measured clinical outcomes, then controlling for them would improve the precision of estimated intervention effects. Thus, each model included the average outcome rate during the baseline period for each pharmacy cluster as an additional regressor. This variable served as a proxy for the net effect of geographic factors on the outcome of interest.

***Independent Variable Construction.*** The indicators for patients' baseline disease status were created in the same way as the indicators for patients' clinical outcomes. The appendix to this chapter shows the relevant ICD-9 codes. The indicators for patients' drug utilization were based on Medicaid drug claims during the baseline period. For all drugs except corticosteroids, the indicator was positive based on a single prescription claim; for corticosteroid, the indicator was positive if the patient used the drug for six months or throughout the baseline. This reflects the nature of the clinical issue under study.

### Models and Statistical Methods

The fact that pharmacy clusters rather than individual patients were randomized also meant that the standard regression framework required some adjustment. If indeed there were factors correlated with cluster that affected outcomes and if, despite the controls included in the models, these factors were not fully measured, then, because the number of clusters was small (38) relative to the number of patients, these differences might not average out between intervention and control groups, despite the large number of patients, leading to misleading estimates of intervention effects. Specifically, traditionally-estimated standard errors would be smaller than the true standard errors.

As has been discussed, the first line of defense against this problem was the effort to include a full set of covariates, in particular the cluster averages discussed above. In addition, a cluster adjusted version of White's method was used to estimate the variance-covariance matrices of the models. This method relaxes the assumption of independence within clusters, and is considered to be a robust variance estimator in cases of cluster-based sampling.[41] An alternative approach would have been to bootstrap the coefficients of each model and derive nonparametric standard errors for the coefficients.[42] For the main population, this alternative method was tested; each anti-depressant model was bootstrapped and standard errors were computed. The results, which are not presented here, were very similar to the results derived using the cluster-adjusted version of White's method.[43]

---

[41] William Rogers originally developed this method in the early 1980's to account for cluster-based sampling issues in the RAND Health Insurance Experiment.

[42] The bootstrapping method involves pulling repeated samples (with replacement), estimating coefficients repeatedly (100 times in this case), and then deriving standard errors from the distributions of estimated coefficients. This method is often used in cases where we have no formulas to compute standard errors or when the assumptions we need to make to compute standard errors do not make sense (i.e. in this case, we could not assume that error terms were independently distributed because of inter-cluster correlation).

[43] Two other methods were considered to account for the effects of clustered sampling. The random effects probit was considered, but the large size of the clusters made this model very difficult if not impossible to implement computationally. Generalized Estimating Equations (GEE) methods could have been used, but this approach would have made estimation more complex and the advantages of this approach were uncertain.

Thus, the regression specification was:

$$y^{c}_{ijt_f} = \alpha + \beta_1 tr_{ij} + X_{it_b}\beta_2 + W^{c}_{it_b}\beta_3 + Z_i\beta_4 + \beta_5 y^{c}_{jt_b} + \nu_{ij}$$

with:

| | |
|---|---|
| $y^{c}_{ijtf}$ | indicator that patient i associated with pharmacy cluster j experienced clinical outcome c during the follow-up period (tf) |
| $tr_{ij}$ | indicator that patient i filled a prescription at treatment pharmacies and not at control pharmacies during the demonstration period |
| $X_{it}$ | general predisposing (socio-demographic) and need characteristics of patient i determined during the baseline period |
| $W^{c}_{itb}$ | patient i's risk factors for outcome c determined during the baseline period |
| $Z_i$ | length of follow-up and baseline windows for patient i |
| $y^{c}_{jtb}$ | rate of clinical outcome c in pharmacy cluster j during the baseline period |
| $\nu_{ij}$ | error |

Because all dependent variables were binary, probit regressions were used. The probit coefficients were transformed to marginal effects to make interpretation easy.[44]

Exhibit 4.9.4 below summarizes all of the models that were estimated in this study. For all of the outcomes except new hip fractures, models were estimated using variables created from all Medicare claims and variables created from inpatient Medicare claims only. The new hip fracture models are based on inpatient claims only. Each model was estimated using the three populations described previously: (1) the main population of dually eligible beneficiaries; (2) a sub-population of dually eligible beneficiaries who used antidepressant drugs in the baseline or follow-up periods; and (3) a sub-population of dually eligible beneficiaries who used antidepressant drugs in the baseline period only. Only models based on all claims and the main population are presented in full in Appendix C (these models are shaded in Exhibit 4.9.4).

---

[44] The marginal effect can be interpreted as the change in the probability of the outcome that is associated with an infinitesimal change in the independent variable, holding other covariates constant. For binary independent variables, the marginal effects are calculated based on a one unit change in the independent variable rather than an infinitesimal change.

Exhibit 4.9.4

## HCFA DUR Demonstration Evaluation Models Estimated

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 1 | Conduction Disorders | Probit | All dual eligibles | 7805 | Based on all Medicare claims. |
| 2 | Conduction Disorders | Probit | All dual eligibles | 7805 | Based on inpatient Medicare claims only. |
| 3 | Conduction Disorders | Probit | Antidepressant drug users in baseline or follow-up | 2121 | Based on all Medicare claims. |
| 4 | Conduction Disorders | Probit | Antidepressant drug users in baseline or follow-up | 2121 | Based on inpatient Medicare claims only. |
| 5 | Conduction Disorders | Probit | Antidepressant drug users in baseline | 1391 | Based on all Medicare claims. |
| 6 | Conduction Disorders | Probit | Antidepressant drug users in baseline | 1391 | Based on inpatient Medicare claims only. |
| 7 | Cardiac Dysrthymia | Probit | All dual eligibles | 7805 | Based on all Medicare claims. |
| 8 | Cardiac Dysrthymia | Probit | All dual eligibles | 7805 | Based on inpatient Medicare claims only. |
| 9 | Cardiac Dysrthymia | Probit | Antidepressant drug users in baseline or follow-up | 2121 | Based on all Medicare claims. |
| 10 | Cardiac Dysrthymia | Probit | Antidepressant drug users in baseline or follow-up | 2121 | Based on inpatient Medicare claims only. |
| 11 | Cardiac Dysrthymia | Probit | Antidepressant drug users in baseline | 1391 | Based on all Medicare claims. |

**Exhibit 4.9.4**

**HCFA DUR Demonstration Evaluation**
**Models Estimated**

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 12 | Cardiac Dysrthymia | Probit | Antidepressant drug users in baseline | 1391 | Based on inpatient Medicare claims only. |
| 13 | Hypotension | Probit | All dual eligibles | 7805 | Based on all Medicare claims. |
| 14 | Hypotension | Probit | All dual eligibles | 7805 | Based on inpatient Medicare claims only. |
| 15 | Hypotension | Probit | Antidepressant drug users in baseline or follow-up | 2121 | Based on all Medicare claims. |
| 16 | Hypotension | Probit | Antidepressant drug users in baseline or follow-up | 2121 | Based on inpatient Medicare claims only. |
| 17 | Hypotension | Probit | Antidepressant drug users in baseline | 1391 | Based on all Medicare claims. |
| 18 | Hypotension | Probit | Antidepressant drug users in baseline | 1391 | Based on inpatient Medicare claims only. |
| 19 | New Hip Fracture | Probit | All dual eligibles | 7805 | Based on inpatient Medicare claims only. |
| 20 | New Hip Fracture | Probit | Antidepressant drug users in baseline or follow-up | 2121 | Based on inpatient Medicare claims only. |
| 21 | New Hip Fracture | Probit | Antidepressant drug users in baseline | 1391 | Based on inpatient Medicare claims only. |
| 22 | Anticholinergic Effects | Probit | All dual eligibles | 7805 | Based on all Medicare claims. |

**Exhibit 4.9.4**

**HCFA DUR Demonstration Evaluation**
**Models Estimated**

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 23 | Anticholinergic Effects | Probit | All dual eligibles | 7805 | Based on inpatient Medicare claims only. |
| 24 | Anticholinergic Effects | Probit | Antidepressant drug users in baseline or follow-up | 2121 | Based on all Medicare claims. |
| 25 | Anticholinergic Effects | Probit | Antidepressant drug users in baseline or follow-up | 2121 | Based on inpatient Medicare claims only. |
| 26 | Anticholinergic Effects | Probit | Antidepressant drug users in baseline | 1391 | Based on all Medicare claims. |
| 27 | Anticholinergic Effects | Probit | Antidepressant drug users in baseline | 1391 | Based on inpatient Medicare claims only. |

*Sources:* International Classification of Diseases, 9th Revision, DHHS Publication No. (PHS) 91-1260, October, 1991.

## Findings

### *Descriptive Statistics*

Exhibit 4.9.5 below displays the means and standard deviations for the dependent variables and the basic set of covariates used in all of the models described above. The treatment group had a larger proportion of non-white individuals relative to the comparison group. Additionally, the treatment group had a statistically significant, lower rate of hypotension compared to the comparison group. Overall, however, the descriptive statistics reveal no striking evidence of a treatment effect; the treatment and comparison groups were quite similar along both baseline characteristics and follow-up outcomes.

Exhibit 4.9.5

### HCFA DUR Demonstration Evaluation
### Means and Standard Deviations of Dependent Variables and Covariates in Main Washington Population Models
### (N = 7805)

| | Treatment (N = 4577) | Control (N = 3228) |
|---|---|---|
| Dependent Variables | | |
| conduction disorders | .035<br>(.184) | .037<br>(.189) |
| cardiac dysrhythmia | .200<br>(.400) | .203<br>(.402) |
| hypotension* | .031<br>(.174) | .039<br>(.194) |
| new hip fracture | .017<br>(.127) | .015<br>(.124) |
| anti-cholinergic effects | .125<br>(.331) | .127<br>(.333) |
| Main Covariates | | |
| female | .736<br>(.441) | .734<br>(.442) |
| white*** | .667<br>(.471) | .732<br>(.443) |
| black*** | .070<br>(.256) | .030<br>(.171) |
| other race*** | .262<br>(.440) | .237<br>(.426) |
| age 65-69*** | .262<br>(.440) | .236<br>(.425) |
| age 70-74 | .208<br>(.406) | .210<br>(.408) |
| age 75-79 | .148<br>(.355) | .160<br>(.367) |
| age 80-84 | .115<br>(.318) | .114<br>(.319) |
| age 85 or higher | .110<br>(.313) | .121<br>(.327) |
| length of baseline period*** | 10.9<br>(2.52) | 10.7<br>(2.70) |
| length of follow-up period*** | 13.3<br>(6.30) | 12.7<br>(6.48) |

Source: Washington Project C.A.R.E. Medicare/Medicaid Files
Note: *** statistically significant difference at the .01 level
** statistically significant difference at the .05 level
* statistically significant difference at the .10 level

## Results

### *Treatment Effects*

Cognitive service payments, by encouraging pharmacist consultations with patients and prescribers, might prevent cardiac problems found to be associated with anti-depressant drug use. It was hypothesized, therefore, that Project C.A.R.E. would be associated with lower rates of conduction disorders and cardiac dysrhythmia, two important adverse effects associated with anti-depressant use. The effect of the intervention was expected to be stronger among those with preexisting cardiac conditions because these individuals are at higher risk for cardiac problems related to antidepressant use.

The estimation results, however, do not suggest that Project C.A.R.E. was effective in preventing conduction disorders or cardiac dysrhythmia. In the conduction disorders models, the estimated treatment effect coefficient is generally inconsistent in sign and not statistically significant across all populations and models. The same is true for the estimated treatment effect coefficient in the cardiac dysrhythmia models.

Project C.A.R.E. also was expected to reduce the rate of hypotension, the most common adverse effect associated with anti-depressant drug use. The model results do suggest that , holding differences between the treatment and control groups constant, there is a small, marginally statistically significant treatment effect for this outcome. In the main population, using outcome indicators based on all claims, Project C.A.R.E. was associated with a less than .01 reduction in the probability of hypotension during the follow-up period. The treatment effect was somewhat larger in magnitude in the two sub-populations of anti-depressant users; for both baseline users as well as baseline and follow-up users, Project C.A.R.E. was associated with a .02 reduction in the probability of hypotension during the follow-up period. The treatment effect was still negative but no longer statistically significant in models based on variables that were created using inpatient claims only.

If Project C.A.R.E was effective in reducing the onset of hypotension, one would expect that the intervention also would be associated with a decrease in new hip fractures, since hip fractures often are a marker of hypotension. The results, however, showed no statistically significant association between Project C.A.R.E. and new hip fractures in any of the different models and populations. This result, therefore, does not support the idea that Project C.A.R.E. was effective in reducing the rate of hypotension.[45]

Finally, it was hypothesized that Project C.A.R.E. would reduce the rate of anticholinergic effects. Pharmacists might discuss patients' potential or existing anticholinergic symptoms with prescribing physicians, or make suggestions directly to patients. These actions in turn could lead to changes in prescriptions and patient behavior that improve anticholinergic symptoms. The models, however, offer no evidence that Project C:A.R.E. was effective in this regard. In the anticholinergic effects models based on variables created using all claims, the estimated treatment effect was never statistically significant. Surprisingly, in the anticholinergic outcomes models based on variables created using inpatient claims only, the treatment effect was positive and it bordered on statistical significance at the .05 level. It is difficult to interpret this result; if the intervention had no effect on anticholinergic effects in general, it is not clear why it would have a positive impact on anticholinergic effects that required inpatient care.

---

[45] The analysis, however, may have missed a large number of hip fracture cases because patients in nursing homes, who might be at high risk for hip fractures, were excluded from the analysis samples.

**Exhibit 4.9.6**

**HCFA DUR Demonstration Evaluation**
**Summary of Signs of Treatment Effect Coefficients**
**Probit Models Based on Variables Created Using All Medicare Claims**

| **Population** | **Dependent Variable** | | | | |
|---|---|---|---|---|---|
| | **Conduction Disorders** | **Cardiac Dysrhythmia** | **Hypotension** | **New Hip Fracture** | **Anti-Cholinergic Effects** |
| Main | -.015<br>(.061) | .015<br>(.037) | -.104*<br>(.057) | .056<br>(.076) | .007<br>(.039) |
| Baseline or Follow-up Users | .143<br>(.117) | .022<br>(.069) | -.191**<br>(.091) | .024<br>(.137) | .059<br>(.070) |
| Baseline Users Only | .169<br>(.151) | .003<br>(.086) | -.202*<br>(.112) | .079<br>(.163) | .118<br>(.088) |

Note:*** difference between arms statistically significant at the 1 percent level
** difference between arms statistically significant at the 5 percent level
* difference between arms statistically significant at the 10 percent level

*Sources:* Iowa Medicare claims

### *Other Covariates*

The signs and statistical significance of the sociodemographic variables' coefficients generally were intuitively appealing. Females were less likely than males to experience cardiac problems, hypotension, and anticholinergic effects, and more likely than males to have a hip fracture in the follow-up period, holding other factors constant. Compared to whites and blacks, individuals of other races were less likely to have all of the outcomes during the follow-up period. Unsurprisingly, individuals who were 85 years old or older were more likely than younger individuals to experience all outcomes. Nevertheless, only the hip fracture model showed evidence of a consistent, direct relationship between age and the outcome.

The best predictor of each outcome in the follow-up period was an indicator for the same outcome during the baseline period. Other baseline disease covariates, and baseline drug utilization markers for diseases, also were good predictors of the outcomes of interest. Cardiac conditions and use of digoxin during the baseline period, for example, were statistically significant predictors of conduction disorders and cardiac dysrhythmia. Baseline use of tricyclic antidepressants and baseline use of anti-hypertensives were both associated with hypotension during the follow-up period.

Across all of the models, the length of the baseline period was negatively associated with the outcomes while the length of the follow-up period was positively associated with the outcomes. Individuals with longer baseline periods perhaps were healthier compared to those with shorter baseline periods, who may have recently left nursing homes. Additionally, individuals with longer baseline periods had better controls for baseline characteristics compared to those with shorter baseline periods. It is not surprising that longer follow-up periods are associated with an increased probability of outcomes because individuals with longer follow-up periods had more time under observation during which an outcome could have occurred.

It was expected that the covariate that captured the wave of the first contact pharmacy would be a statistically significant predictor of outcomes if the efficiency and the effectiveness of cognitive services provision improved over time. Pharmacies that entered the demonstration in later waves, for example, might have learned from the experiences of pharmacies that entered during earlier waves; better provision of cognitive services, in turn, may have led to improved outcomes. The estimated coefficients on the wave variables, however, were generally not statistically significant. This finding suggests that a learning effect either did not exist or that the effect was not strong enough to measurably affect outcomes.

### *Pharmacy Cluster Effects*

To account for the fact that clusters of pharmacies rather than individuals were randomized to the treatment and control groups, each model included as a covariate the average rate of the dependent variable in the individual's pharmacy cluster. It was expected that this variable would help control for unobservable, cluster level differences in factors that affect the outcomes of interest. These variables did turn out to be statistically significant and positive in sign in several models, but, generally, the variables did not display a consistent pattern in sign across the models and they were often statistically insignificant. The Huber/White variance estimator, which also was used to adjust for cluster-based sampling, led to standard errors that were very similar to those calculated using a conventional variance estimate. Overall, then, adjustment for cluster-based sampling did not have any impact on the interpretation of the findings.

### Comment

Project C.A.R.E. was intended to encourage the provision of cognitive services by reimbursing pharmacists for these activities. The increase in cognitive services, in turn, was expected to reduce the rate of adverse clinical outcomes that result from inappropriate prescription drug use. This study focused on the impact of Project C.A.R.E. on a set of clinical outcomes related to antidepressant drug use: (1) conduction disorders and dysrhythmia; (2) hypotension; and (3) anticholinergic effects. The analysis controlled for treatment and comparison group characteristics that may have differed at baseline, as well as the cluster-based sampling design used in the demonstration.

The findings indicate that Project C.A.R.E had no measurable impact on the outcomes of interest. Although some models suggested that the intervention was associated with a small decrease in hypotension, the validity of this finding is suspect because of the large number of hypothesis tests conducted.[46] Evidence from the demonstration supports the idea that payments for cognitive services did increase the rate of cognitive services provision, particularly later in the demonstration. These results, therefore, suggest that even though the rate of cognitive services increased, the increase was not effective in reducing adverse clinical outcomes related to antidepressant use.

There are a number of reasons why this effect might not exist, or why the effect might be too small to detect using this type of analysis. Project C.A.R.E. was not focused specifically on preventing adverse outcomes related to antidepressant use; instead, the intervention was designed to target Medicaid users of any prescription drug. Furthermore, about half of cognitive services interventions that were performed focused on patient related problems, such as compliance. Only about 28 percent of cognitive services events resulted in a change in drug therapy [Christensen (1996)]. It is possible, then, that only a very small proportion of cognitive services events resulted in changes that even had the potential to prevent an adverse outcome related to antidepressant use. This is particularly true for rare outcomes. Project C.A.R.E. therefore may not have effected measurable improvements in the rate of adverse effects related to antidepressant drugs.

It is more likely, however, that limitations in the study design clouded measurement of an existing treatment effect, particularly if the treatment effect was very small. Generally, a cluster sampling approach reduces the power of detecting differences between the treatment and control groups. In this case, however, the intracluster correlation is likely to be low; cluster based sampling therefore probably does not seriously reduce the power of the study design.

More important is the fact that the treatment effect is expected to be very small — it is hypothesized that the intervention might lead to half a percentage point difference in outcomes between the treatment and control groups. To capture such tiny differences, huge sample sizes would be needed. With the sample sizes used in the main population, the study design had between 7 and 30 percent power to detect a difference of this size in outcomes.[47] The limited power of the study therefore may have had some impact on the findings.

## 4.10 The Impact of the Cognitive Services Demonstration on Adverse Clinical Outcomes Related to the Use of Benzodiazepines

Drug utilization review is a process that has the goal of ensuring rational drug therapy through attention to prescribing, dispensing, and use of pharmaceuticals [Knapp and colleagues, (1989a,1989b); Yates, Rupp, and Schondelmeyer (1990); Byrns 1993)]. In the context of prescription drug use, rational therapy means providing drug therapy targeted to optimal therapeutic outcomes while being cost-effective [AMA (1991); Lipton and Bird (1991)].

---

[46] Although a formal Boneferroni correction was not implemented, it is reasonable to expect that one would reject the null hypothesis when the null hypothesis was true in at least one case. The marginal statistical significance of the treatment coefficients in the hypotension models, and the inconsistent results that were obtained from the inpatient claims based hypotension models cast further doubt on the finding of a significant, negative association between Project C.A.R.E. and hypotension.

[47] The sample sizes required to detect a half percentage point difference in antidepressant outcomes range from about 11,000 to over 100,000.

In practice, DUR is a dynamic and variable process that is directed to one or more of three target behaviors: inappropriate prescribing by the physician, inappropriate dispensing by the pharmacist, and inappropriate use by the patient. DUR may be conducted on various levels depending on the focus and intensity of efforts. At the lowest level DUR might be as simple as retrospective review of dispensing and payment data to ascertain contract compliance and surveillance for fraud and abuse. At a much higher level DUR can involve application of computer algorithms to patient level data (current drug use, diagnoses, etc.) at the point of sale (dispensing) to identify both patients at risk and prescribers who manifest irrational prescribing practices [Kralewski et al. (1990)0. Basic to the concept of point-of-sale DUR is the notion that the pharmacist is included in the decision path. Pharmacists are trained to deal with the varying sensitivity and specificity of computer-assisted DUR algorithyms. Further, pharmacists can, if necessary, proactively provide appropriate counseling to both patients and their prescribers to reduce risks of adverse consequences. There is a widely held belief that these cognitive services, as provided by pharmacists, can be an effective mechanism to promote rational drug therapy and, at the same time, reduce unnecessary expenditures.

The Project C.A.R.E. program was a demonstration that offered payments to pharmacists for providing cognitive services. Payment for cognitive service was intended to provide appropriate financial compensation for the time and effort needed to provide cognitive services. The demonstration was conducted in Washington State from February 1994 to September 1995. The population of interest in this study was Washington residents who were eligible for Medicare and Medicaid, who were over age 65, not enrolled in an HMO, not residing in a nursing home, and who filled a prescription at a demonstration pharmacy at least once. This chapter assessed the effects of the C.A.R.E. program on the occurrence of falls and hip fractures, adverse, clinical outcomes that have been linked to inappropriate prescribing of benzodiazepines. As this chapter will show, holding other factors constant, treatment and control group patients generally did not have statistically significant differences in these outcomes. The analysis presented in this chapter therefore indicates that the Washington C.A.R.E. project was not effective in preventing adverse outcomes associated with inappropriate prescribing of benzodiazepines.

### Adverse Clinical Outcomes Related to Use of Benzodiazepines

The elderly are well represented in the population of benzodiazepines users. Although the elderly account for only about 13% of the U.S. population they account for 35 to 40% of the prescriptions written for hypnotic agents [Gottlieb (1990)]. Clinical research has strongly suggested that the use and misuse of benzodiazepines is linked to adverse clinical effects among the elderly [Markowitz (1996)]. The following sections provide a description of the outcomes selected for this study[48] and a brief review of the evidence for their association with benzodiazepine use.

---

[48] In the planning phase for this study the rate of motor vehicle accidents had been identified as a clinical outcome that would be included in the benzodiazepine analyses. However, preliminary examination of the occurrence of moter vehicle accidents as encoded in the Medicare files indicated that there were too few cases (main population, N = 30) to warrant further analyses.

### *Falls*

Falls and their sequelae are a major public health problem among the elderly. Studies on the incidence of falling among the elderly suggest that on average, approximately 45 percent of community-dwelling persons over the age of 65 will fall at least once each year [Cummings and Klineberg (1994); O'Loughlin, et al. (1993); Campbell, Borrie, & Spears (1989)]. Falls are the leading accident that occurs among the elderly and falls are the leading cause of death from injury [Rice and MacKenzie (1989)]. Hospital admissions from falls increase with aging reaching 4% per annum for men and 7% per annum for women age 85 and above [Lord (1990)].

Numerous studies support the conclusion that use of psychotropic drugs, especially benzodiazepines, is associated with risk for falling [Cumming (1998); Monane and Avorn (1996); Wickham et al. (1989); Sorock & Shimkin (1988)]. Weiner, Hanlon and Studenski (1998) conducted a prospective study on community-dwelling male veterans. They found that the risk of falls was significantly greater in central nervous system-active medication users as compared to non-users and fall liability increased with multiple central nervous system-active medication use. Herings, Stricker, deBoer, Bakker, and Sturmans (1995) conducted a case-control study of persons age 55 and above, to examine the risk factor associated with falling leading to femur fractures. They found that falling was significantly associated with use of benzodiazepines, use of multiple benzodiazepines, and particularly with use of short half-life benzodiazepines. They concluded that the strong dose-response relationship with both short and long half-life benzodiazepines argues that the relationship between benzodiazepines with falls and femur fractures is probably due to high dosages rather than elimination half-life.

### *Hip Fractures*

The evidence for a link between psychotropic medications and hip fractures is mixed. A recent review of the epidemiology of hip fractures is provided by Cumming, Nevitt, and Cummings (1997). Two large population-based case control studies by Ray and his colleagues [Ray et al. (1987,1989, 1991)] examined the association of medications with hip fractures. The earliest of these studies [Ray et al. (1987)] involved linkage of records of prescription drug use and medical diagnoses of Medicaid beneficiaries. They found that antipsychotics, long-acting benzodiazepines, and tricyclics antidepressants were all significantly associated with increased risk for hip fracture. These finding were confirmed in the second study using claims data from the health insurance system in Saskatchewan, Canada [Ray (1989, 1991)]. The second study also found that opioid analgesics (codeine and/or propoxyphene) were also associated with hip fractures.

## Evaluation Approach

The objective of this study was to examine the association between the demonstration factor — payment for cognitive services — and the occurrence of selected adverse events related to use of benzodiazepines. The expectation was that payment for cognitive services would increase the number of cognitive service interventions provided. This increase, in turn, was expected to change prescriber and/or patient behavior, and make prescriptions — especially prescriptions with known risks such as psychotropics — safer and more effective. The occurrence of adverse clinical outcomes related to prescribing of benzodiazepines was therefore expected to decrease as a result of the safer prescribing and use of this class of prescription drugs.

The analytic strategy of this study was multivariate. For each candidate clinical outcome, the residual effect of the demonstration, after adjustment for known risk factors was examined. Multivariate regression procedures were used to control for demographic, disease, and drug factors predisposing a given patient to the outcome under study. All control variables were measured during a baseline period prior to the onset of the demonstration so as to be independent of any demonstration effects.

### Data

The data used in this study was extracted from Medicaid drug claims and Medicare health services utilization claims for patients in the state of Washington in 1993 (to provide a baseline) and 1994-1995. (The demonstration took place between February of 1994 and September of 1995.) Other records and data streams generated by the demonstration were also incorporated into the data set. Data development is described in detail in Section 4.7.

### Definitions of Baseline and Follow-up Periods

Pharmacies were recruited into the demonstration in three distinct waves (February, April, and September of 1994)[49]. This study assumes that a patient's exposure to the demonstration (and consequently the patient's follow-up period) began when a patient first filled a prescription at a pharmacy participating in the demonstration. In the majority of cases, the follow-up period then continued until the demonstration ended in September 1995.

All persons in the study were ascertained to be eligible for Medicaid and Medicare for the full period of the demonstration. In certain cases, if the patient lost eligibility for Medicare or Medicaid, entered a nursing home, or entered an HMO, the follow-up period ended sooner because the outcomes data captured in the claims would be incomplete. A patient's death was also deemed to terminate the follow-up period.

Each patient's baseline periods might also differ in length. The baseline period was determined to end the day before the patient's first contact with a demonstration pharmacy and to reach back twelve months or until the patient first became eligible for the Medicare program, became eligible for the Medicaid program, left an HMO, or left a nursing home and became resident in the community. Again, the latter three restrictions are imposed in order to ensure that the medical record is complete during the baseline period.

Note that clinical outcomes that occurred on the day the patient first went to a demonstration pharmacy are excluded from both the baseline and the follow-up periods. In such cases, it is impossible to know whether the clinical issue precipitated the prescription or the prescription led to the clinical issue. The conservative decision therefore was to omit the outcome from both periods.

---

[49] In principle, effects of the intervention might last beyond the demonstration period. However, limited payment for CS was made available to both intervention and control pharmacies after September of 1995 which could confound any estimated intervention effects based on outcomes after that date.

## Study Populations

As displayed below in Exhibit 4.10.1, the data were used to create three study populations: (1) a main study population; (2) a sub-sample of the main study population consisting of those who had been using benzodiazepines during the baseline period; and (3) a sub-sample of the main population consisting of those who had been using benzodiazepines during the baseline or the follow-up period.[50] Although all models were implemented using these three samples, not all results are presented in this chapter. Instead, this chapter focuses on interpreting and summarizing the results from the analyses; selected models estimated using the main population are presented in full in the appendix to this chapter.

**Exhibit 4.10.1**

**HCFA DUR Demonstration Evaluation**
**Study Populations**

| Study Population: | Description: |
|---|---|
| Main Population<br>(N = 7804) | Dually eligible for Medicare and Medicaid, at least 65 years old, not enrolled in an HMO or living in a nursing home, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies |
| Baseline or Follow-up Benzodiazepines Users<br>(N = 1703) | Dually eligible for Medicare and Medicaid, at least 65 years old, not enrolled in an HMO or living in a nursing home, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled at least one prescription for a benzodiazepine during the baseline period or during the follow-up period |
| Baseline Benzodiazepine Users<br>(N = 1000) | Dually eligible for Medicare and Medicaid, at least 65 years old, not enrolled in an HMO or living in a nursing home, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled at least one prescription for a benzodiazepine during the baseline period |

Sources: Washington Project C.A.R.E. Medicare/Medicaid Files

---

[50] If Project C.A.R.E. affected prescribing of non-benzodiazepines (e.g., trials of buspirone) in the follow-up period and the population of interest consists of benzodiazepines users in the baseline and follow-up periods, a selection problem would exist. In this case, it is possible that those who benefitted from the intervention would be more likely than others to be excluded from the sample.

The main study population consisted of elderly patients (65 and over on January 1, 1993) who were eligible for both the Medicaid and Medicare programs, who were not enrolled in an HMO during the study period, who were not living in a nursing home, and who filled a prescription at either intervention or control pharmacies during the demonstration period. As noted above, the elderly are particularly at risk for adverse outcomes related to benzodiazepine use and misprescribing. Thus, the elderly are hypothesized to be more likely than other people of other age groups to benefit from the intervention. The requirement of dual eligibility and exclusion if enrolled in an HMO ensures that the analyses have complete and consistent records of drug and health services utilization for all patients in the data set.

In addition to the main population two sub-samples were defined: (1) individuals who used benzodiazepines during the baseline period, and (2) individuals who used benzodiazepines during the baseline or follow-up periods. Narrowing the population in this way was intended to improve the specificity of the analysis. Furthermore, the analysis considered both a sub-sample of baseline benzodiazepine users as well as a sub-sample of baseline or follow-up benzodiazepine users because of concerns that the intervention might affect benzodiazepine use in the follow-up period. In this case, estimates based on the baseline/follow-up benzodiazepine drug sub-sample might be biased.

### Dependent Variables

The outcomes selected for this study were falls and hip fractures. Outcome indicators were constructed using Medicare diagnosis codes.[51] Each outcome was associated with a series of ICD-9 codes. For falls, if a patient's record showed evidence of one of those ICD-9s as either the primary or secondary diagnosis for an inpatient stay or for a physician encounter (paid by Part A or Part B of Medicare) during the follow-up period, then the patient was determined to have experienced the event. A model also was estimated in which the falls variable was based on inpatient claims alone. The rationale for this is based on the possibility of differential validity of the outcome diagnoses. It might be possible that compared to physicians in the outpatient setting, hospitals might be more precise in their coding of diagnoses because of its payment implications. Variables that were created using inpatient claims only therefore might contain less measurement error compared to variables created using all claims.

For hip fractures, indicator variables were computed using records of inpatient stays and physician encounters and from inpatient stays only. Further, an indicator variable based on verification of hip fractures via CPT codes was computed. For the main population, there were 103 hip fractures observed using the CPT information, 130 events using inpatient claims, and 187 hip fractures recorded using all Medicare claims. Preliminary analyses for the main population showed good correspondence between the indicator variable based on inpatient records and the CPT verified events. It was felt that the higher rate of hip fractures observed using all Medicare claims probably exists because outpatient, follow-up care, for old hip fractures has been coded as new hip fractures. For this study, new hip fractures based on inpatient claims was selected as the appropriate outcomes event. Exhibit 4.10.2 presents the associated ICD-9 codes used to define the indicator variables for these outcomes.

---

[51] Because all of the samples used in this analysis included individuals who were dually eligible for Medicare and Medicaid, it is assumed that all medical services were initially billed and primarily covered by Medicare.

**Exhibit 4.10.2**

**HCFA DUR Demonstration Evaluation**
**C.A.R.E. Evaluation on Benzodiazepine Outcomes, Dependent Variables**

| Variable: | ICD-9 Classification: |
|---|---|
| Fall Occurence | E880-E888 |
| New Hip Fractures | 820.0-820.9, confirmed by CPT Codes |

*Sources:* International Classification of Diseases, 9th Revision, DHHS Publication No. (PHS) 91-1260, October, 1991.

## Independent Variables

### *Basic Covariates*

Each model included a basic set of covariates that captured exposure to the intervention as well as general predisposing and need characteristics. Additionally, covariates were added to each model to control for clinical risk factors that are specific to each of the outcomes. The main set of independent variables, as well as the outcome-specific independent variables, are described in the following sections.

## Exposure to the Intervention

Although this demonstration was designed as a random experiment, randomization occurred at the level of the pharmacy cluster while clinical benefits of the demonstration were expected to accrue at the level of the patient.[52] This posed a significant methodological problem. Individuals had the potential to fill prescriptions at different pharmacies and, potentially, to fill prescriptions at a mix of intervention, control, and non-demonstration pharmacies. It could be expected that the effect of the demonstration would be diluted for such individuals depending on the relative proportions of prescriptions filled at the different pharmacy types. To resolve this problem, this study adopted the following approach. Individuals were considered members of the intervention group if they ever had a Medicaid-paid drug claim at an intervention pharmacy and never had such a claim at a control pharmacy. Similarly, they were considered members of the control group if they ever had a Medicaid-paid drug claim at a control pharmacy and never had such a claim at an intervention pharmacy. Individuals who went to both intervention and control pharmacies were excluded from the study regardless of the number of filled prescription at either intervention or control pharmacies[53].

---

[52] Pharmacy clusters, built around common prescribers, were created and randomized to minimize the influence that intervention pharmacies might exert through common prescribers on control patients. In Washington there were 38 pharmacy clusters ranging in size from 1 person to 1200 people.

[53] For this purpose, the day that the patient first went to the demonstration pharmacy was included in the follow-up period.

### General Predisposing and Need Characteristics

The basic analytic strategy for the several clinical outcomes studies reported here was to examine the effect of the treatment variable within a regression framework that statistically controls for demographic, drug use, and disease factors. For all outcomes, the regression models included variables for the patient's race, sex, and age. These factors affect health outcomes indirectly as proxies for health status. They may also affect the probability of having access to, or using, health care services. Thus these measures serve to control for general predisposing and need characteristics. In this study on benzodiazepine-related clinical outcomes, age, race, and sex are also known to act directly as risk factors to the outcomes of interest.

### Length of Baseline and Follow-up Periods

It would be expected that patients with a longer follow-up period would be more likely to experience any given clinical outcome than one with a shorter follow-up period. Therefore, the length of the follow-up period was included as a covariate in the models. While the length of the baseline period does not affect patients' probabilities of experiencing a specific clinical problem, it might be expected that the assessment of health spending, number of drugs, and risk factors will tend to be measured with less precision for patients with short baseline periods relative to patients with full baseline periods. Thus, controlling for other measured covariates, the probability of clinical outcomes might be higher for patients with short baselines; consequently, the length of the baseline period was included as an additional covariate. The lengths of the follow-up and baseline period were entered in models as the natural log of the number of months in each period.

### Pharmacy Wave

Pharmacies entered the cognitive services demonstration in three waves. The first wave of 160 pharmacies enrolled in February 1994. A second wave (N = 33) and a third wave (N = 7) of pharmacies entered in April and September 1994. It is possible that pharmacies entering the demonstration later learned from the experiences of prior entrants and provided more effective cognitive services. Additionally, the rate of cognitive service events per 100 dispensed prescriptions varied widely over time in the treatment group. For these reasons, indicator variables were included in the models that account for the wave of the first demonstration pharmacy where the individual filled a prescription. The reference category in the models was wave one. The basic set of independent variables that were used in every model (as described above) are summarized below in Exhibit 4.10.3.

**Exhibit 4.10.3**

**HCFA DUR Demonstration Evaluation**
**Basic Set of Covariates Included in All Models**

race
sex
age (in categories)
length of baseline period
length of follow-up period
wave of first contact demonstration pharmacy

Source: Washington Project C.A.R.E. Medicare/Medicaid Files

### Covariate Risk Factors Specific to Each Outcome

In addition to the basic set of independent variables each regression model included measures of risk factors specific to each outcome.

#### *Falls*

There are many risk factors for falls. Risk factors for falls generally include the conditions or characteristics of persons, including physiological impairments, acute or chronic diseases, and medications that impair sensory, cognitive, or musculoskeletal functioning and which increase the risk of falls [Cumming (1998)]. The evidence for multiple concurrent prescription drug use is fairly strong [Myers et al. (1991); Cumming et al. (1991)] . Many studies have identified psychotropic medications, especially antidepressants, to be risk factors for falls [Cumming (1998)]. There is also evidence implicating selected cardiovascular, analgesic, and other classes of medications as risk factors for falls through their direct and indirect effects on psychomotor functioning and balance [Cumming (1998); Monane and Avorn (1996)]. Since the number of specific drugs that have been implicated for falls is large, and also because the primary focus of the study is on the effect of the intervention rather than drug use per se, it was deemed appropriate to reduce the number of covariates by developing indicator variables based on clusters of medication use. For the falls outcome, three clusters were defined, psychotropics and related, cardiovascular and related, and other risk-related medications. For multiple prescription drugs an indicator covariate was constructed to indicate whether or not an individual was concomitantly taking medications representing three or more different therapeutic classes. Exhibit 4.10.4 presents the set of risk factors included in analyses involving falls.

#### *Hip Fractures*

Numerous studies have identified risk factors for hip fractures. Diseases, conditions, and medications that affect bone density have been repeatedly identified as risk factors. Additionally, those diseases, conditions, and medications that predispose to falls are known to be associated with risk for hip fractures. As indicated above, most studies have identified multiple medications to be positively associated with falls. For hip fractures, on the other hand, there is some evidence that thiazide diuretics may be protective by reducing calcium excretion. For this reason, this drug group was retained as a single independent covariate whereas the other medications known to be associated with hip fracture risk were aggregated as described above for the falls outcome. Exhibit 4.10.5 present the covariates included in models that examine the effect of the intervention on the hip fracture outcome.

**Exhibit 4.10.4**

## HCFA DUR Demonstration Evaluation
## C.A.R.E. Evaluation
## Disease, Condition, and Drug Use Covariates for Models Involving Falls as the Dependent Variable

- Disease/Condition Covariates
  - Arthritis
  - Convulsions
  - Dementia
  - Depression
  - Diabetes
  - Epilepsy
  - Falls
  - Gait Abnormality
  - Hip Fracture
  - Hypotension
  - Orthostatic hypotension
  - Syncope and Collapse

- Medication Use Covariates
  - Psychotropics and related medications
    - Antidepressants
    - Antipsychotics
    - Anticonvulsants
    - Non-benzodiazepine CNS depressants

  - Cardiovascular and related medications
    - Antiarrythmics
    - Beta blockers
    - Calcium channel blockers
    - Digoxin

  - Other medications related to risk of falls
    - Aminoglycosides
    - Antihistamines
    - Narcotic analgesics

  - Multiple prescriptions

Source: Washington Project C.A.R.E. Medicare/Medicaid Files

**Exhibit 4.10.5**

## HCFA DUR Demonstration Evaluation

## C.A.R.E. Evaluation

## Disease, Condition, and Drug Use Covariates for Models Involving Hip Fractures as the Dependent Variable

- Disease/Condition Covariates
    - Arthritis
    - Dementia
    - Diabetes
    - Epilepsy
    - Falls
    - Gait Abnormality
    - Hip Fracture
    - Hypotension
    - Orthostatic hypotension
    - Osteoporosis
    - Osteomalacia
    - Syncope and Collapse

- Medication Use/Covariates
    - Psychotropics and related
        - Antidepressants
        - Antipsychotics
        - Anticonvulsants
        - Non-benzodiazepine CNS depressants

    - Cardiovascular and related
        - Antiarrythmics
        - Beta Blockers
        - Calcium Channel Blockers
        - Digoxin

    - Other medications related to increased risk of hip fracture
    - Aminoglycosides
        - Antihistamines
        - Corticosteroids
        - Estrogens
        - Narcotic analgesics

    - Thiazide diuretics

    - Multiple prescriptions

Source: Washington Project C.A.R.E. Medicare/Medicaid Files

### Rate of Clinical Outcome in Pharmacy Cluster

The pharmacy cluster, the unit of randomization in this analysis, effectively defined medical communities. Potentially, there could be unmeasured differences in demographic and behavioral characteristics among patients in different pharmacy clusters that influenced the need for care and the predisposition to seek it. Similarly, there could be geographic variation associated with cluster in access to care and in its practice, such as variations in coding of conditions and diagnoses and in treatment protocols. If these differences were associated with measured clinical outcomes, then controlling for them would improve the precision of estimated intervention effects. Thus, each model included the average outcome rate during the baseline period for each pharmacy cluster as an additional regressor. This variable served as a proxy for the net effect of geographic factors on the outcome of interest.

### Independent Variable Construction

The indicators for patient's baseline disease and drug use status were created in the same way as the indicators for patients' clinical outcomes. The indicators for diseases and conditions were based on diagnosis codes for Medicare claims during the baseline period. The prior hip fracture indicator was based on inpatient claims only. The indicators for patients' drug utilization were based on Medicaid drug claims during the baseline period. For all drugs, the indicator was positive based on one or more claims represented in the claims history.

## Models and Statistical Methods

Because all dependent variables were binary, probit regressions were used. The regression specification for the examination of treatment effects in this study was:

$$y^{c}_{ijt_f} = \alpha + \beta_1 tr_{ij} + X_{it_b}\beta_2 + W^{c}_{it_b}\beta_3 + Z_i\beta_4 + \beta_5 y^{c}_{jt_b} + v_{ij}$$

with:

| | |
|---|---|
| $y^c_{ijtf}$ | indicator that patient i associated with pharmacy cluster j experienced clinical outcome c during the follow-up period (tf) |
| $tr_{ij}$ | indicator that patient i filled a prescription at treatment pharmacies and not at control pharmacies during the demonstration period |
| $X_{it}$ | general predisposing (socio-demographic) and need characteristics of patient i determined during the baseline period |
| $W^c_{itb}$ | patient i's risk factors for outcome c determined during the baseline period |
| $Z_i$ | length of follow-up and baseline windows for patient i |
| $y^c_{jtb}$ | rate of clinical outcome c in pharmacy cluster j during the baseline period |
| $v_{ij}$ | error |

Exhibit 4.10.6 below summarizes all of the models that were estimated in this study. For falls, models were estimated using variables created from all Medicare claims and variables created from inpatient Medicare claims only. The new hip fracture models are based on inpatient claims only. Each model was estimated using the three populations described previously: (1) the main population of dually eligible beneficiaries; (2) a sub-population of dually eligible beneficiaries who used benzodiazepines in the baseline or follow-up periods; and (3) a sub-population of dually eligible beneficiaries who used benzodiazepines in the baseline period only. Only models based on all claims and the main population are presented in full in Appendix D (these models are shaded in Exhibit 4.10.6).

**Exhibit 4.10.6**

**HCFA DUR Demonstration Evaluation**
**Models Estimated**

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 1 | Falls | Probit | All dual eligibles | 7804 | Based on all Medicare claims. |
| 2 | Falls | Probit | All dual eligibles | 7804 | Based on inpatient Medicare claims only. |
| 3 | Falls | Probit | Benzodiazepine users in baseline | 1000 | Based on all Medicare claims. |
| 4 | Falls | Probit | Benzodiazepine users in baseline | 1000 | Based on inpatient Medicare claims only. |
| 5 | Falls | Probit | Benzodiazepine users in baseline or follow-up | 1703 | Based on all Medicare claims. |
| 6 | Falls | Probit | Benzodiazepine users in baseline or follow-up | 1703 | Based on inpatient Medicare claims only. |
| 7 | Hip Fractures | Probit | All dual eligibles | 7804 | Based on inpatient Medicare claims only. |
| 8 | Hip Fractures | Probit | Benzodiazepine users in baseline | 1000 | Based on inpatient Medicare claims only. |
| 9 | Hip Fractures | Probit | Benzodiazepine users in baseline or follow-up | 1703 | Based on inpatient Medicare claims only. |

## Findings

### *Descriptive Statistics*

Exhibit 4.10.7 below displays the means and standard deviations for the dependent variables and the basic set of covariates used in all of the models described above. The treatment group had a larger proportion of non-white individuals relative to the comparison group. Further examination of the race by treatment contingency tables showed that there were significant differences in the occurrence of the outcome variables according to race but that these differences were not associated with the treatment effect. For example, while blacks and others were significantly less likely to experience falls compared to whites, there were no significant differences in the proportions of falls between the treatment and control areas for either blacks or others. Further, while blacks and other non-white groups were over represented in the treatment areas, there were no occurrences of hip fractures among blacks in the follow-up period in either the treatment or control areas[54].

Examination of Exhibit 4.10.7 indicates other differences between treatment and controls areas. The treatment group had higher representation of young elderly compared to the control group. There was a significant but small differences in the lengths of the baseline and follow-up periods between the treatment and control groups. Overall, the comparison of mean differences reveals no striking evidence of a treatment effect for the outcome measures; the treatment and comparison groups were quite similar.

### *Treatment Effects*

It was hypothesized that Project C.A.R.E. would lead to increased rates of cognitive services interventions, and that these cognitive services interventions, in turn, would lead to reduced risk of adverse outcomes related to benzodiazepine use. This analysis focused on two preventable adverse outcomes related to benzodiazepines use: (1) falls; and (2) hip fractures. If the increase in cognitive services did significantly decrease rates of these problems, the models should indicate that, after adjusting for baseline demographic and risk factors and for factors related to potential design effects, the treatment group should have had lower rates of adverse benzodiazepine effects relative to the control group. If this were the case, the probit coefficients should be negative and significant.

---

[54] As a consequence of the lack of variability in the outcome measures, the solutions for the probit models for hip fractures were numerically unstable. In order to include race as a covariate it was necessary to create an indicator variable (nonwhite) that included blacks and others.

**Exhibit 4.10.7**

### HCFA DUR Demonstration Evaluation
### Means and Standard Deviations of Dependent Variables and Covariates in Main Washington Population Models
### (N = 7805)

| | Treatment (N = 4577) | Control (N = 3227) |
|---|---|---|
| Dependent Variables | | |
| falls | .0551 | .0536 |
| (inpatient/outpatient) | (.2253) | (.2253) |
| falls | .0159 | .0170 |
| (inpatient) | (.1253) | (.1295) |
| new hip fracture | .0155 | .0149 |
| (inpatient) | (.1211) | (1236) |
| Main Covariates | | |
| female | .7363 | .7344 |
| | (.4417) | (.4417) |
| white*** | .6675 | .7319 |
| | (.4712) | (.4430) |
| black*** | .0704 | .0301 |
| | (.2558) | (.1708) |
| other race*** | .2622 | .2380 |
| | (.4399) | (.4259) |
| age 65-69*** | .3218 | .2863 |
| | (.4672) | (.4521) |
| age 70-74 | .2495 | .2575 |
| | (.4328) | (.4373) |
| age 75-79 | .1813 | .1955 |
| | (.3853) | (.3967) |
| age 80-84 | .1376 | .1391 |
| | (.3446) | (.3461) |
| age 85 or higher | .1097 | .1215 |
| | (.3125) | (.3267) |
| length of baseline | 10.935 | 10.736 |
| period*** | (2.520) | (.2.701 |
| length of follow-up | 13.270 | 12.691 |
| period*** | (6.298) | (6.480) |
| wave 1 pharmacy*** | .7743 | .7961 |
| | (.4181) | (4030) |
| wave 2 pharmacy*** | .1835 | .1571 |
| | (3871) | (.3640) |
| wave 3 pharmacy | .0422 | .0468 |
| | (.2001) | (.2112) |

Source: Washington Project C.A.R.E. Medicare/Medicaid Files

Note *** statistically significant difference at the .01 level
** statistically significant difference at the .05 level
* statistically significant difference at the .10 level

Exhibit 4.10.9 in the Appendix provides the coefficients and their respective standard errors for all variables included in the analysis of falls based on the main population. Analogous results for new hip fractures in the follow-up period are provided in Exhibit 4.10.10 of the Appendix. The probit results for the treatment effects for all nine models are summarized in Exhibit 4.10.8.

**Exhibit 4.10.8**

**HCFA DUR Demonstration Evaluation**
**Washington C.A.R.E. Demonstration Evaluation**
**Summary of Treatment Effects: Probit Coefficients**

| Population | Dependent Variable | | |
|---|---|---|---|
| | Falls, all Medicare records | Falls, inpatient Medicare records only | Hip fractures, inpatient Medicare records only |
| Main | .030<br>(.045) | .001<br>(.075) | .062<br>(.079) |
| Baseline Benzodiazepine Users Only | -.050<br>(.130) | -.107<br>(.185) | -.147<br>(.215) |
| Baseline and Follow-up Benzodiazepine Users Only | .080<br>(.099) | .006<br>(.144) | .090<br>(.153) |

Note:*** difference between arms statistically significant at the 1 percent level
** difference between arms statistically significant at the 5 percent level
* difference between arms statistically significant at the 10 percent level

Source: Washington Project C.A.R.E. Medicare/Medicaid Files

Examination of the estimation results summarized in Exhibit 4.10.8 suggests that Project C.A.R.E. was not effective in preventing either of fall or hip fracture rates in the demonstration area relative to the control area. None of the coefficients approach conventional levels of significance. While the direction of the observed effects over the several sub-populations is generally consistent with expectations, the magnitude of the effects is very low and non-significant.

### *Other Covariates*

The results relating to risk factors for falls and hip fractures were generally as expected. Age is positively associated with risk of falling and of experiencing a hip fractures. The occurrence of falls is positively associated with the length of the follow-up period. Prior history of falls during the baseline period is significantly predictive of future falls during the follow-up period.

Preliminary bivariate analyses — mentioned above — had indicated that, compared to blacks and others, whites had a greater risk of hip fractures. This was partially confirmed in the multivariate analyses; for the main population, white persons had a significantly greater risk of hip fractures. Although the effect was in the expected direction for the two other sub-samples the magnitude was not large enough to be statistically significant.

Multiple concomitant prescription drug use–sometimes also called polypharmacy–was not significantly associated with hip fracture risk. This may be, in part, due to to the inclusion of the other separate therapeutic class risk categories. On the other hand, this indicator was significantly associated with increased risk of falls as measured by inpatient records and nearly significant ($p < .13$) for falls based on inpatient and outpatient Medicare claims.

Use of psychotropic medications was consistently positively associated with hip fracture risk but the magnitude of the effect was not significant. None of the indicator drug use indicator variables was associated with falls. However, use of thiazide diuretics appears to be protective against hip fractures as is indicated by recent studies. All three models showed evidence of a protective effect for hip fractures but the effect size only reached conventional significance levels ($p < .05$) for the subsample of person who were also using benzodiazepines in the baseline period.

### *Pharmacy Cluster Effects*

To account for the fact that clusters of pharmacies rather than individuals were randomized to the treatment and control groups, each model included as a covariate the average rate of the dependent variable in the individual's pharmacy cluster. It was expected that this variable would help control for unobservable, cluster level differences in factors that affect the outcomes of interest. These variables did indeed turn out to be statistically significant in three of the six models relating to falls. Nevertheless, the direction of the effect was inconsistent in sign. and magnitude of the marginal effect of the cluster rates on the outcomes was usually very close to zero. Overall, the adjustment for cluster-based sampling did not have any impact on the interpretation of the findings.

## Comment

The Washington C.A.R.E. demonstration was designed to provide an incentive for pharmacists to detect and deal with prescription problems at the point of service so that the risk for adverse outcomes could be reduced or avoided. This chapter focused on the impact of the intervention on two clinical outcomes related to benzodiazepine use: falls and hip fractures. The analysis controlled for treatment and comparison group characteristics that may have differed at baseline, as well as the cluster-based sampling design used in the demonstration. Based on the regression results discussed above, there is no evidence that the cognitive services related to benzodiazepine therapy prevented either of the two adverse outcomes that expected to be responsive to the intervention.

There are several reasons that may offered to explain why no significant effect this effect was found. It may be that the risk situations that were detected were not able to be greatly ameliorated due to the importance of continuing the clinical care for which the drugs were prescribed.

Further, even if changes in drug therapy were possible, in order for a reduction in outcomes to occur, a specific chain of risk-reducing actions must take place once a prescription has been submitted to the pharmacist. These actions usually require a coordinated set of responses on the part of the pharmacist, the prescriber and the patient. The failure of the cognitive service intervention to produce observable reductions in falls and fractures may be a reflection of a systematic failure of one or more of the responsible parties to accept and act on the problematic therapy.

It is also possible that limitations in the study design complicated the measurement and detection of existing treatment effects, particularly if the treatment effects were very small. Generally, a cluster sampling approach reduces the power of detecting differences between the treatment and control groups. In this case, however, the intracluster correlation is likely to be low, and many of the clusters were very small; cluster based sampling therefore does not seriously reduce the power of the study design.

Another important factor is that the treatment effect is expected to be very small. With the sample sizes used and the low observed incidence of falls and fractures in the sub-population of baseline and follow-up period benzodiazepine drug users, the study had somewhat limited power to detect small differences between treatment and control groups. To assess small differences between treatment and control groups huge sample sizes would be needed. Although the sample size constraints may have limited the ability of the study to detect significant effects, the findings nevertheless suggest that the Washington C.A.R.E. demonstration did not lead to large enough effects to be observed without multivariate analyses. Moreover, attempts to detect even a small effect via statistical control of many risk factors were unsuccessful.

## 4.11 The Impact of Project C.A.R.E. on Adverse Outcomes Related to Inappropriate Prescribing of Cardiovascular Drugs

Pharmacists play a critical role in ensuring that patients receive and use medications appropriately. In addition to product selection, packaging, labeling, and basic patient counseling, pharmacists may interact with patients and/or prescribers to uncover a potential or existing prescribing problem, or to suggest a more effective, less toxic, or less costly drug regimen [Christensen et.al. (1996)]. As the link between the prescriber and the patient, the pharmacist is in a unique position to provide these types of enhanced, *cognitive* services.

Cognitive services provided by pharmacists can prevent a range of prescribing problems, including drug-drug interactions, drug-disease interactions, dosage errors, duplicative therapy, and patient noncompliance. The vulnerable populations served by Medicaid, which include indigent elderly people and children, are at higher risk for these prescribing problems and their associated adverse outcomes [Christensen et.al. (1996)]. These groups therefore are the most likely to benefit from an effective cognitive services program.

Project C.A.R.E. was a demonstration that offered payments to pharmacists for providing cognitive services. The demonstration was conducted in Washington State from February 1994 to September 1995. This study assessed the effects of the C.A.R.E. program on a set of adverse, clinical outcomes that have been linked to inappropriate prescribing of cardiovascular drugs. The population of interest in this study was elderly, Washington residents who were eligible for Medicare and Medicaid, and who filled a prescription at a demonstration pharmacy at least once.

The findings presented in this chapter indicate that Project C.A.R.E. was *not* associated with a reduction in adverse outcomes related to inappropriate prescribing of cardiovascular drugs. Holding baseline characteristics constant, individuals who visited pharmacies where pharmacists were paid for cognitive services provision were just as likely as control group patients to experience hyperkalemia, cardiac conditions, and symptoms and diagnoses of digoxin toxicity. The analysis indicates, therefore, that the intervention did not improve cardiovascular drug therapy to the extent that major, adverse health outcomes were prevented.

## Adverse Clinical Outcomes Related to Use of Cardiovascular Drugs

This study focuses on the use of three cardiovascular drugs: ACE inhibitors, verapamil, and digoxin. Clinical research has indicated that the use of these three drugs is associated with a number of adverse outcomes, including hyperkalemia, cardiovascular problems, and digoxin toxicity. In the following sections, the relationships of interest in this study are motivated by a brief literature review of recent clinical research in this area.

### *ACE Inhibitors and Hyperkalemia*

The relationship between use of ACE inhibitors and hyperkalemia, a condition that is characterized by excessive amounts of potassium in the blood, is of particular interest in this study. Among patients with renal insufficiency, the body's potassium balance is maintained through an increase in excretion in the remaining nephrons of the kidney. This compensatory action is mediated by an increase in aldosterone. ACE inhibitors can suppress the alodosterone produced by the adrenal gland, causing potassium to accumulate [Schlueter et. al. (1994)]. Potassium also can accumulate if ACE inhibitors are prescribed with potassium or potassium sparing diuretics.

Hyperkalemia is is more common in patients with chronic renal failure, diabetes, autonomic neuropathy, and adrenal insufficiency [Howes (1995)]. Risk groups for hyperkalemia also include users of certain drugs, including salt substitutes, some potassium sparing diuretics, ACE inhibitors, NSAIDS, beta-adrenergic antagonists, cyclosporine, heparin, and digoxin. Although patients with advanced renal failure frequently have mild hyperkalemia, life threatening hyperkalemia is uncommon until the glomerular filtration rate (GFR) falls below 5-10 mL per minute. In almost all cases, hyperkalemia can be controlled by taking patients off the offending drugs and/or by administering other medications such as cation exchange resin, insulin and glucose [Rakel (1996)].

In a hospital study conducted through the Vermont College of Medicine, 308 patients (1.7 percent of all patients admitted over the course of one year) were identified as having severe hyperkalemia, and of these 172 were found to have sustained hyperkalemia. The authors reported that 43 (14.0 percent) of these sustained cases were found to be temporally related to a drug or drugs that could cause hyperkalemia. The most common drug associated with hyperkalemia in this population was potassium chloride, but catopril was implicated in about 20 percent of cases. Many of the patients with severe hyperkalemia had predisposing conditions, mostly renal insufficiency and diabetes [Rimmer et. al. (1987)].

The results of this study do not provide a clear sense of either incidence or prevalence, and they cannot be generalized to settings outside the hospital. Nevertheless, the results do document that hyperkalemia is linked to use of ACE inhibitors with some frequency, especially in certain high-risk groups.

### *Verapamil and Cardiovascular Events*

Verapamil is a calcium channel blocker used to manage hypertension and to prevent and control angina pectoris. The action of verapamil is as a vasodilator; it also acts as a depressant to myocardial contractility and the caridiac conduction system, particularly at the AV node. Calcium channel blockers as a class have been shown to be effective in lowering blood pressure, in relieving symptoms in patients with stable angina, and in increasing excerise tolerance, but not necessarily in preventing serious morbidity and mortality [Silvestry and Kimmel (1996)].

There has been concern that verapamil may increase the risk of adverse, cardiovascular outcomes, particularly when used with beta blockers or with other calcium channel blockers, or when used at certain dosage levels. A number of studies have focused on morbidity and mortality associated with verapamil following an acute myocardial infarction. The Danish Verapamil Infarction Trials (DAVIT) I and II were randomized, double blind, placebo-controlled trials with 717 treatment group patients (who received verapamil) and 719 controls (who received placebo) in the first trial, and 878 treatment group patients and 897 controls in the second trial [The Danish Study Group on Verapamil in Myocardial Infarction (1990)].

In the first DAVIT trial, reinfarction and mortality rates were lower in the treatment group than the control group, but this difference was not statistically significant. The second DAVIT trial showed another non-significant difference between the treatment and control groups in mortality rates, but there was a statistically significant reduction in first major event rates at 18 months in the treatment group. Sub-group analyses indicated that treatment group patients without heart failure had statistically significant reductions in mortality, cardiac events, and renal infarction compared to control group patients without heart failure. In a sub-group of people *with* existing heart failure, there were no statistically significant differences between the treatment and control groups.

When verapamil is used in combination with a beta blocker, it is possible that the inotropic, chronotropic, and dromotropic effects of the two drugs will be additive. This effect could result in am increased probability of conduction defects and heart failure [Sever (1989)]. Opie (1988), in his description of the addition of verapamil to a regimen that already includes propranolol or metoprolol, cites a number of studies that inidcate that this combination may lead to adverse cardiovascular outcomes. These adverse cardiovascular outcomes include reduced contractility [Packer et. al. (1982)], increased heart size [Johnstone et. al. (1985)], symptomatic sinus bradycardia [Winniford et. al. (1985)], and depressed left ventricular function [Johnstone et. al. (1985)]. Peargien and Benowitz (1991) note that when verapamil is given in combination with beta blockers, bradycardia, hypotension, and conduction disturbances may occur, especially in the presence of underlying heart disease [citing Hamann et. al. (1985, 1987)].

Researchers have found that verapamil use is associated with hypotension. Pearigen and Benowitz (1991) report that out of a series of 91 patients who were identified through a regional poison control center for overdose of calcium channel blockers, 9 out of the 38 patients were hypotensive [citing Ramoska et. al. (1990)]. In a review of drugs that can cause orthostatic hypotension, Schoenberger (1991) classifies calcium channel blockers as infrequent in provoking this outcome, while dopamine agonists, for example, were considered to cause this outcome frequently, and for tricyclic antidepressants, hypotension was a common problem. Schoenberger (1991) cites MacMahon (1990) as reporting that hypotension may occur in approximately 5 percent of patients taking dihydropyridine calcium channel blockers, and in less than 2.5 percent of patients in the other classes of calcium channel blockers.

### *Digoxin and Digoxin Toxicity*

Digoxin has been a major component of therapy for controlling heart rate during atrial fibrilation. Digoxin has a narrow therapeutic window, however, and toxicity is a risk. Furthermore, digoxin can interact with other drugs that may be used to treat atrial fibrilation, such as verapamil, quinidine, and procainamide. Interaction with verapamil may reduce clearance of digoxin, resulting in an increase of blood digoxin levels of up to 60-90 percent [Opie (1988), Opie citing Pedersen (1985)]. Digoxin toxicity is most common among the elderly.

## Evaluation Approach

The objective of this analysis was to explore the relationship between payment for cognitive services and the rate of adverse events related to use of cardiovascular drugs. It was hypothesized that payment for cognitive services would increase the number of cognitive service interventions provided. This increase, in turn, was expected to change prescriber and/or patient behavior, and make prescriptons safer and more effective. The rate of adverse outcomes related to prescribing of cardiovascular drugs was therefore expected to decrease, reflecting improvements in prescribing and patient compliance. The general objective of this study was to compare rates of clinical outcomes among patients who went to intervention pharmacies to those among patients who went to control pharmacies. Multiple regression was used in order to control for demographic, disease, and drug factors predisposing a given patient to the outcome under study. All control variables were measured during a baseline period before the demonstration began so as to be independent of any demonstration effects. A version of White's method that accounts for cluster based sampling was used to estimate standard errors [Rogers (1993)].

## Data

The data used in this study was extracted from Medicaid drug claims and Medicare health services utilization claims for patients in the state of Washington in 1993 (to provide a baseline) and 1994-1995. (The demonstration took place between February of 1994 and September of 1995.) Other records and data streams generated by the demonstration were also incorporated into the data set. Data development is described in detail in Chapter 4.7.

### *Study Populations*

As displayed below in Exhibit 4.11.1, the data was used to create three study populations: (1) a main study population; (2) a sub-sample of the main study population consisting of those who had been using ACE inhibitors, calcium channel blockers, and/or digoxin during the baseline period; and (3) a sub-sample of the main stduy population consisting of those who had been using ACE inhibitors, calcium channel blockers, and/or digoxin during the baseline or the follow-up period. Although all models were implemented using these three samples, not all results are presented in this chapter. Instead, this chapter focuses on interpreting and summarizing the results from these models; selected models estimated using the main population are presented in full in the appendix to this chapter.

The main study population consisted of elderly patients (65 and over on January 1, 1993) who were eligible for both the Medicaid and Medicare programs and who went to either intervention or control pharmacies during the intervention. The elderly are particularly at risk for adverse outcomes related to

**Exhibit 4.11.1**

**HCFA DUR Demonstration Evaluation**
**Study Populations**

| Study Population: | Description: |
|---|---|
| Main Population<br>(N = 7805) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies |
| Baseline or Follow-up Cardiac Drug Users<br>(N = 4251) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled at least one prescription for ACE inhibitors, calcium channel blockers, and/or digoxin during the baseline period or during the follow-up period |
| Baseline Cardiac Drug Users<br>(N = 3138) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled at least one prescription for ACE inhibitors, calcium channel blockers, and/or digoxin during the baseline period |

*Sources:* WA Medicaid claims

cardiovascular drug use. Moreover, they are generally in more delicate health and taking more drugs than younger patients; the elderly thus are hypothesized to be more likely than other people of other age groups to benefit from the intervention. The requirement of dual eligibility ensures that the analyses have complete and consistent records of drug and health services utilization for all patients in the data set.

The study population was further limited to individuals who used cardiac drugs during the baseline period, and individuals who used cardiac drugs during the baseline or follow-up periods. Narrowing the population in this way was intended to improve the specificity of the analysis.[55] Furthermore, the analysis considered both a sub-sample of baseline cardiac drug users as well as a sub-sample of baseline or follow-up cardiac drug users because of concerns that the intervention might affect cardiac drug use in the follow-up period. In this case, estimates based on the baseline/follow-up cardiac drug sub-sample might be biased.[56]

---

[55] Nevertheless, analyses based on this more narrow population would miss any intervention effect that involved patients avoiding the use of cardiovascular drugs altogether.

[56] If Project C.A.R.E. affected cardiac drug use in the follow-up period and the population of interest consists of cardiac drug users in the baseline and follow-up periods, a selection problem would exist. In this case, it's possible that those who benefitted from the intervention would be more likely than other to be excluded from the sample.

### *Definitions of Baseline and Follow-up Periods*

Pharmacies enrolled in the demonstration in three waves (February, April, and September of 1994). Although other approaches would be possible, this study assumes that a patient's exposure to the demonstration (and consequently the patient's follow-up period) began when a patient first filled a prescription at a pharmacy participating in the demonstration. In the majority of cases, the follow-up period then continued until the demonstration ended in September 1995.[57] In certain cases, if the patient lost eligibility for Medicare or Medicaid, entered a nursing home, or entered an HMO, the follow-up period ended sooner because the outcomes data captured in the claims would be incomplete after that event.[58] A patient's death was also deemed to terminate the follow-up period.

In a similar way, patients' baseline periods differed in length. The baseline period was determined to end the day before the patient's first contact with a demonstration pharmacy and to reach back twelve months or until the patient first became eligible for the Medicare program, became eligible for the Medicaid program, left an HMO, or left a nursing home and became resident in the community. Again, the latter three restrictions are imposed in order to ensure that the medical record is complete during the baseline period.

Note that clinical outcomes that occurred on the day the patient first went to a demonstration pharmacy are excluded from both the baseline and the follow-up periods. In such cases, it is impossible to know whether the clinical issue precipitated the prescription or the prescription led to the clinical issue. The safe decision therefore was to omit the outcome from both periods.

### *Dependent Variables*

Based on the research literature, reviewed briefly above, three outcomes were selected for this study: (1) hyperkalemia, related to the use of ACE inhibitors; (2) congestive heart failure and other adverse cardiac events related to the use of verapamil, a calcium channel blocker; and (3) diagnosis and symptoms of digoxin toxicity, related to the use of digoxin. To capture cardiac problems related to verapamil use, two measures were constructed; one measure included diagnosis of congestive heart failure only, while the other, broader measure included congestive heart failure, conduction disorders, sinus bradycardia and hypotension. Digoxin toxicity was measured using four different variables that ranged from a specific diagnosis of digoxin toxicity to a broad range of symptoms that are associated with digoxin toxicity. These dependent variables are summarized below in Exhibit 4.11.2.

---

[57] In principle, effects of the intervention might last beyond the demonstration period. However, limited payment for CS was made available to both intervention and control pharmacies after September of 1995 which could confound any estimated intervention effects based on outcomes after that date.

[58] Nursing home prescriptions that were batched and sent to demonstration pharmacies logically could not be affected by cognitive services interventions.

**Exhibit 4.11.2**

**HCFA DUR Demonstration Evaluation Dependent Variables**

| Variable: | ICD-9 Classification: |
|---|---|
| hyperkalemia | 276.7 |
| congestive heart failure (CHF) | 428.0 |
| congestive heart failure or conduction disorders or hypotension or sinus bradycardia (CHF+) | 428.0, all 426, 458.0-458.9, 427.81-427.89 |
| any one of digoxin toxicity diagnosis, atrial ventricular block, conduction disorders, sinus bradycardia, ventricular tachycardia, abdominal pain, anorexia, disturbed vision, fatigue, headache or nausea/vomiting (all symptoms and diagnosis) | 972.1, 426.00-426.13, all 426, 427.81-427.89, 427.1, 789.00-789.09, 783.0, 368.9, 780.7, 784.0, 787.01-787.03 |
| digoxin toxicity diagnosis only | 972.1 |
| arrhythmia complex only (any one of atrial ventricular block, conduction disorders, sinus bradycardia, or ventricular tachycardia) | 426.00-426.13, all 426, 427.81 -427.89, 427.1 |
| symptom complex only (any one of abdominal pain, anorexia, disturbed vision, fatigue, headache, or nausea/vomiting) | 789.00-789.09, 783.0, 368.9, 780.7, 784.0, 787.01-787.03 |

*Sources:* International Classification of Diseases, 9th Revision, DHHS Publication No. (PHS) 91-1260, October, 1991.

Outcome indicators were constructed using Medicare diagnosis codes.[59] Each outcome was associated with a series of ICD-9 codes. If a patient's record showed evidence of one of those ICD-9s as either the primary or secondary diagnosis for an inpatient stay or for a physician encounter (paid by Part A or Part B of Medicare) during the follow-up period, then the patient was determined to have had the outcome.

For each outcome, a model also was estimated in which the outcome variable was based on inpatient claims alone. It is possible that compared to physicians in the outpatient setting, hospitals might be more precise in their coding of diagnoses because of its payment implications. Variables that were created using inpatient claims only therefore might contain less measurement error compared to variables created using all claims. Models based on variables created using inpatient claims only are not presented here because the results were very similar to models based on variables created using all claims.

---

[59] Because all of the samples used in this analysis included individuals who were dually eligible for Medicare and Medicaid, it is assumed that all medical services were initially billed and primarily covered by Medicare.

### *Independent Variables*

Each model included a basic set of covariates that captured exposure to the intervention as well as general predisposing and need characteristics. Additionally, covariates were added to each model to control for clinical risk factors that are specific to each of the outcomes. The main set of independent variables, as well as the outcome specific independent variables, are described in the following sections.

***Exposure to the Intervention.*** Although this demonstration was designed as a random experiment, randomization occurred at the level of the pharmacy cluster while clinical benefits accrued at the level of the patient.[60] This posed several methodological challenges. First, individuals had the potential to go to multiple pharmacies and, potentially, to go to a mix of intervention, control, and non-demonstration pharmacies. The study adopted a simple approach. Individuals were considered members of the intervention group if they ever had a Medicaid-paid drug claim at an intervention pharmacy and never had such a claim at a control pharmacy. They were considered members of the control group if they ever had a Medicaid-paid drug claim at a control pharmacy and never had such a claim at an intervention pharmacy. Individuals who went to both intervention and control pharmacies were excluded from the study.[61]

***General Predisposing and Need Characteristics.*** The basic design of the study was to compare treatment and control patients within a regression framework that controls for demographic, drug, and disease factors. For all five outcomes, regressions included controls for the patient's race, sex, and age. These factors affect health outcomes both directly and as proxies for behavior; they may also affect the predisposition to use care. These variables, which were taken from the enrollment records of the Medicaid program, were included in regressions as a series of indicator variables.

### Length of Baseline and Follow-up Periods

Other things equal, a patient with a longer follow-up period would be more likely to experience any clinical outcome than one with a shorter follow-up period, thus the length of the follow-up period was made an additional control in models. While the length of the baseline period does not affect patients' probabilities of clinical problems in and of itself, expected health spending, number of drugs, and risk factors will tend to be slightly under measured for patients with short baselines relative to patients with full baselines. Thus, controlling for measured covariates, the probability of clinical outcomes might be higher for patients with short baselines; consequently, the length of the baseline period was included as an additional control. The lengths of the follow-up and baseline period were entered in models as the natural log of the number of months in each period.

---

[60] Pharmacy clusters, built around common prescribers, were created and randomized to minimize the influence that intervention pharmacies might exert through common prescribers on control patients. In WA, there were 38 pharmacy clusters ranging in size from 1 person to 1200 people.

[61] For this purpose, the day that the patient first went to the demonstration pharmacy was included in the follow-up period.

*Pharmacy Wave*

Pharmacies entered the cognitive services demonstration in three waves. The first wave of 160 pharmacies enrolled in February 1994. A second wave (N = 33) and a third wave (N = 7) of pharmacies entered in April and September 1994. It is possible that pharmacies entering the demonstration later learned from the experiences of prior entrants and provided more effective cognitive services. Additionally, the rate of cognitive service events per 100 dispensed prescriptions varied widely over time in the treatment group. For these reasons, indicator variables were included in the models that capture the wave of the first demonstration pharmacy where the individual filled a prescription. The omitted category in the models was wave one.

The basic set of independent variables that were used in every model (as described above) are summarized below in Exhibit 4.11.3.

**Exhibit 4.11.3**

**HCFA DUR Demonstration Evaluation**
**Basic Set of Covariates Included in All Models**

race
sex
age (in categories)
length of baseline period
length of follow-up period
wave of first contact demonstration pharmacy

Source: Washington Project C.A.R.E. Medicare/Medicaid Files

### Risk Factors Specific to Each Outcome

*Hyperkalemia.* Analyses of hyperkalemia added additional controls for pre-existing hyperkalemia, baseline renal insufficiency, baseline congestive heart failure, and baseline diabetes, as captured by baseline claims for insulin and oral hypoglacemics. These conditions are risk factors for ongoing conduction disorders and cardiac dysrhythmia. The models also controlled for the use of beta blockers in the baseline period because use of these drugs is associated with hyperkalemia.

*Congestive Heart Failure/Conduction Disorders/Hypotension/Sinus Bradycardia.* Analyses of these outcomes controlled for pre-existing congestive heart failure, baseline conduction disorders, baseline hypotension, baseline sinus bradycardia, and other cardiac conditions diagnosed during the baseline period.

*Digoxin Toxicity Diagnosis/Symptom Complex/Arrythmia Complex.* Analyses of these outcomes controlled for baseline digoxin toxicity diagnosis and symptoms, baseline renal insufficiency, baseline congestive heart failure, and baseline diabetes (as captured by claims for insulin and oral hypoglacemics). These conditions are risk factors for digoxin toxicity. The use of certain drugs in combination with digoxin is also associated with digoxin toxicity. For this reason, baseline use of verapamil, quinidine, amiodarone, propafenone, indomethacin, itraconazole, aplrazolam, and diuretics were included as covariates in the digoxin toxicity models.

### Rate of Clinical Outcome in Pharmacy Cluster

The pharmacy cluster, the unit of randomization in this analysis, effectively defined medical communities. Potentially, there could be unmeasured differences in demographic and behavioral characteristics among patients in different pharmacy clusters that influenced the need for care and the predisposition to seek it. Similarly, there could be geographic variation associated with cluster in access to care and in its practice, such as variations in coding of conditions and diagnoses and in treatment protocols. If these differences were associated with measured clinical outcomes, then controlling for them would improve the precision of estimated intervention effects. Thus, each model included the average outcome rate during the baseline period for each pharmacy cluster as an additional regressor. This variable served as a proxy for the net effect of geographic factors on the outcome of interest.

***Independent Variable Construction.*** The indicators for patients' baseline disease status were created in the same way as the indicators for patients' clinical outcomes. The appendix to this chapter shows the relevant ICD-9 codes. The indicators for patients' drug utilization were based on Medicaid drug claims during the baseline period. For all drugs, the indicator was positive based on a single prescription claim.

### Models and Statistical Methods

The fact that pharmacy clusters rather than individual patients were randomized also meant that the standard regression framework required some adjustment. If indeed there were factors correlated with cluster that affected outcomes and if, despite the controls included in the models, these factors were not fully measured, then, because the number of clusters was small (38) relative to the number of patients, these differences might not average out between intervention and control groups, despite the large number of patients, leading to misleading estimates of intervention effects. Specifically, traditionally-estimated standard errors would be smaller than the true standard errors.

As has been discussed, the first line of defense against this problem was the effort to include a full set of covariates, in particular the cluster averages discussed above. In addition, a cluster adjusted version of White's method was used to estimate the variance-covariance matrices of the models. This method relaxes the assumption of independence within clusters, and is considered to be a robust variance estimator in cases of cluster-based sampling.[62] An alternative approach would have been to bootstrap the coefficients of each model and derive nonparametric standard errors for the coefficients.[63] For the main population, this alternative method was tested; each anti-depressant model was bootstrapped and standard errors were computed. The results, which are not presented here, were very similar to the results derived using the cluster-adjusted version of White's method.[64]

---

[62] William Rogers originally developed this method in the early 1980's to account for cluster-based sampling issues in the RAND Health Insurance Experiment.

[63] The bootstrapping method involves pulling repeated samples (with replacement), estimating coefficients repeatedly (100 times in this case), and then deriving standard errors from the distributions of estimated coefficients. This method is often used in cases where we have no formulas to compute standard errors or when the assumptions we need to make to compute standard errors do not make sense (i.e. in this case, we could not assume that error terms were independently distributed because of inter-cluster correlation).

Thus, the regression specification was:

$$y^c_{ijt_f} = \alpha + \beta_1 tr_{ij} + X_{it_b}\beta_2 + W^c_{it_b}\beta_3 + Z_i\beta_4 + \beta_5 y^c_{jt_b} + v_{ij}$$

with:

| | |
|---|---|
| $y^c_{ijtf}$ | indicator that patient i associated with pharmacy cluster j experienced clinical outcome c during the follow-up period (tf) |
| $tr_{ij}$ | indicator that patient i filled a prescription at treatment pharmacies and not at control pharmacies during the demonstration period |
| $X_{it}$ | general predisposing (socio-demographic) and need characteristics of patient i determined during the baseline period |
| $W^c_{itb}$ | patient i's risk factors for outcome c determined during the baseline period |
| $Z_i$ | length of follow-up and baseline windows for patient i |
| $y^c_{jtb}$ | rate of clinical outcome c in pharmacy cluster j during the baseline period |
| $v_{ij}$ | error |

Because all dependent variables were binary, probit regressions were used. The probit coefficients were transformed to marginal effects to make interpretation easy.[65]

Exhibit 4.11.4 below summarizes all of the models that were estimated in this study. For all of the outcomes except new hip fractures, models were estimated using variables created from all Medicare claims and variables created from inpatient Medicare claims only. The new hip fracture models are based on inpatient claims only. Each model was estimated using the three populations described previously: (1) the main population of dually eligible beneficiaries; (2) a sub-population of dually-eligible beneficiaries who used cardiac drugs in the baseline or follow-up periods; and (3) a sub-population of dually eligible beneficiaries who used cardiac drugs in the baseline period only. Only models based on all claims and the main population are presented in full in Appendix E (these models are shaded in Exhibit 4.11.4).

---

[64] Two other methods were considered to account for the effects of clustered sampling. The random effects probit was considered, but the large size of the clusters made this model very difficult if not impossible to implement computationally. Generalized Estimating Equations (GEE) methods could have been used, but this approach would have made estimation more complex and the advantages of this approach were uncertain.

[65] The marginal effect can be interpreted as the change in the probability of the outcome that is associated with an infinitesimal change in the independent variable, holding other covariates constant. For binary independent variables, the marginal effects are calculated based on a one unit change in the independent variable rather than an infinitesimal change.

Exhibit 4.11.4

## HCFA DUR Demonstration Evaluation Models Estimated

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 1 | Hyperkalemia | Probit | All dual eligibles | 7805 | Based on all Medicare claims. |
| 2 | Hyperkalemia | Probit | All dual eligibles | 7805 | Based on inpatient Medicare claims only. |
| 3 | Hyperkalemia | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on all Medicare claims. |
| 4 | Hyperkalemia | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on inpatient Medicare claims only. |
| 5 | Hyperkalemia | Probit | Cardiac drug users in baseline | 3536 | Based on all Medicare claims. |
| 6 | Hyperkalemia | Probit | Cardiac drug users in baseline | 3536 | Based on inpatient Medicare claims only. |
| 7 | Congestive Heart Failure | Probit | All dual eligibles | 7805 | Based on all Medicare claims. |
| 8 | Congestive Heart Failure | Probit | All dual eligibles | 7805 | Based on inpatient Medicare claims only. |
| 9 | Congestive Heart Falure | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on all Medicare claims. |
| 10 | Congestive Heart Failure | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on inpatient Medicare claims only. |
| 11 | Congestive Heart Failure | Probit | Cardiac drug users in baseline | 3536 | Based on all Medicare claims. |

Exhibit 4.11.4

## HCFA DUR Demonstration Evaluation Models Estimated

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 12 | Congestive Heart Failure | Probit | Cardiac drug users in baseline | 3536 | Based on inpatient Medicare claims only. |
| 13 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | All dual eligibles | 7805 | Based on all Medicare claims. |
| 14 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | All dual eligibles | 7805 | Based on inpatient Medicare claims only. |
| 15 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on all Medicare claims. |
| 16 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on inpatient Medicare claims only. |
| 17 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | Cardiac drug users in baseline | 3536 | Based on all Medicare claims. |
| 18 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | Cardiac drug users in baseline | 3536 | Based on inpatient Medicare claims only. |

**Exhibit 4.11.4**

**HCFA DUR Demonstration Evaluation Models Estimated**

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 19 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | All dual eligibles | 7805 | Based on all Medicare claims. |
| 20 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | All dual eligibles | 7805 | Based on inpatient Medicare claims only. |
| 21 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on inpatient Medicare claims only. |
| 22 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on all Medicare claims. |
| 23 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | Cardiac drug users in baseline | 3536 | Based on all Medicare claims. |
| 24 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | Cardiac drug users in baseline | 3536 | Based on inpatient Medicare claims only. |
| 25 | Digoxin Toxicity Diagnosis | Probit | All dual eligibles | 7805 | Based on all Medicare claims. |
| 26 | Digoxin Toxicity Diagnosis | Probit | All dual eligibles | 7805 | Based on inpatient Medicare claims only. |
| 27 | Digoxin Toxicity Diagnosis | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on all Medicare claims. |

**Exhibit 4.11.4**

**HCFA DUR Demonstration Evaluation**
**Models Estimated**

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 28 | Digoxin Toxicity Diagnosis | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on inpatient Medicare claims only. |
| 29 | Digoxin Toxicity Diagnosis | Probit | Cardiac drug users in baseline | 3536 | Based on all Medicare claims. |
| 30 | Digoxin Toxicity Diagnosis | Probit | Cardiac drug users in baseline | 3536 | Based on all Medicare claims. |
| 31 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | All dual eligibles | 7805 | Based on all Medicare claims. |
| 32 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | All dual eligibles | 7805 | Based on inpatient Medicare claims only. |
| 33 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on all Medicare claims. |
| 34 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on inpatient Medicare claims only. |
| 35 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | Cardiac drug users in baseline | 3536 | Based on all Medicare claims. |

**Exhibit 4.11.4**

## HCFA DUR Demonstration Evaluation Models Estimated

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 36 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | Cardiac drug users in baseline | 3636 | Based on inpatient Medicare claims only. |
| 37 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | All dual eligibles | 7805 | Based on all Medicare claims. |
| 38 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | All dual eligibles | 7805 | Based on inpatient Medicare claims only. |
| 39 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on all Medicare claims. |
| 40 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | Cardiac drug users in baseline or follow-up | 4251 | Based on inpatient Medicare claims only. |
| 41 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | Cardiac drug users in baseline | 3536 | Based on all Medicare claims. |
| 42 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | Cardiac drug users in baseline | 3636 | Based on inpatient Medicare claims only. |

*Sources:* International Classification of Diseases, 9th Revision, DHHS Publication No. (PHS) 91-1260, October, 1991.

## Findings

### *Descriptive Statistics*

Exhibit 4.11.5 below displays the means and standard deviations for the dependent variables and the basic set of covariates used in all of the models described above. The treatment group had a larger proportion of non-white individuals relative to the comparison group. Overall, the descriptive statistics reveal no striking evidence of a treatment effect;the treatment and comparison groups were quite similar along both baseline characteristics and follow-up outcomes.

## Results

### *Treatment Effects*

It was hypothesized that Project C.A.R.E. would lead to increased rates of cognitive services interventions, and that these cognitive services interventions, in turn, would lead to reduced risk of adverse outcomes related to cardiovascular drug use. This analysis focused on three, preventable adverse outcomes related to cardiovascular drug use: (1) ACE inhibitors and hyperkalemia; (2) verapamil, a calcium channel blocker, and cardiac problems; and (3) digoxin and digoxin toxicity. If the increase in cognitive services did significantly decrease rates of these problems, the models should indicate that, after adjusting for baseline factors, the treatment group had lower rates of adverse effects relative to the control group.

The estimation results, however, do not suggest that Project C.A.R.E. was effective in preventing any of the outcomes studied. Exhibit 4.11.6 presents the signs of the treatment coefficients across the seven models. Across all of the models estimated, the treatment effect was negative but it was never statistically significant at conventional levels. There is no evidence, therefore, that treatment group patients were less likely than control group patients to suffer from adverse outcomes related to inappropriate prescribing of cardiovascular drugs.

### Other Covariates

Females were less likely than males to have arrhythmia complex, and more likely than males have symptom complex; otherwise, gender was not a significant predictor of most outcomes. Generally, individuals of other races were less likely than whites and blacks to experience adverse outcomes. Although there was no consistent relationship between all of the age categories and the outcomes. the most aged (85+ years) were more likely than younger individuals to experience all outcomes.

Baseline measures of diseases and drug utilization that were expected to place individuals at higher risk for outcomes were included as covariates in the models. These covariates were excellent predictors of the outcomes. Hyperkalemia in the follow-up period was significantly associated with baseline hyperkalemia as well as baseline renal insufficiency, baseline congestive heart failure, and baseline diabetes, as measured by the use of insulin and oral hypoglacemics. Congestive heart failure, hypotension, conduction disorders and sinus bradycardia in the follow-up period were significantly associated with baseline cardiac conditions. Digoxin toxicity was significantly associated with

## Findings

### *Descriptive Statistics*

Exhibit 4.11.5 below displays the means and standard deviations for the dependent variables and the basic set of covariates used in all of the models described above. The treatment group had a larger proportion of non-white individuals relative to the comparison group. Overall, the descriptive statistics reveal no striking evidence of a treatment effect;the treatment and comparison groups were quite similar along both baseline characteristics and follow-up outcomes.

## Results

### *Treatment Effects*

It was hypothesized that Project C.A.R.E. would lead to increased rates of cognitive services interventions, and that these cognitive services interventions, in turn, would lead to reduced risk of adverse outcomes related to cardiovascular drug use. This analysis focused on three, preventable adverse outcomes related to cardiovascular drug use: (1) ACE inhibitors and hyperkalemia; (2) verapamil, a calcium channel blocker, and cardiac problems; and (3) digoxin and digoxin toxicity. If the increase in cognitive services did significantly decrease rates of these problems, the models should indicate that, after adjusting for baseline factors, the treatment group had lower rates of adverse effects relative to the control group.

The estimation results, however, do not suggest that Project C.A.R.E. was effective in preventing any of the outcomes studied. Exhibit 4.11.6 presents the signs of the treatment coefficients across the seven models. Across all of the models estimated, the treatment effect was negative but it was never statistically significant at conventional levels. There is no evidence, therefore, that treatment group patients were less likely than control group patients to suffer from adverse outcomes related to inappropriate prescribing of cardiovascular drugs.

### Other Covariates

Females were less likely than males to have arrhythmia complex, and more likely than males have symptom complex; otherwise, gender was not a significant predictor of most outcomes. Generally, individuals of other races were less likely than whites and blacks to experience adverse outcomes. Although there was no consistent relationship between all of the age categories and the outcomes, the most aged (85+ years) were more likely than younger individuals to experience all outcomes.

Baseline measures of diseases and drug utilization that were expected to place individuals at higher risk for outcomes were included as covariates in the models. These covariates were excellent predictors of the outcomes. Hyperkalemia in the follow-up period was significantly associated with baseline hyperkalemia as well as baseline renal insufficiency, baseline congestive heart failure, and baseline diabetes, as measured by the use of insulin and oral hypoglacemics. Congestive heart failure, hypotension, conduction disorders and sinus bradycardia in the follow-up period were significantly associated with baseline cardiac conditions. Digoxin toxicity was significantly associated with

Exhibit 4.11.5

**HCFA DUR Demonstration Evaluation**
**Means and Standard Deviations of Dependent Variables and Covariates in Main Washington Population Models (N = 7805)**

| | Treatment (N = 4577) | Control (N = 3228) |
|---|---|---|
| Dependent Variables | | |
| hyperkalemia | .016<br>(.002) | .018<br>(.002) |
| congestive heart failure | .236<br>(.006) | .235<br>(.007) |
| congestive heart failure, conduction disorders, hypotension, or sinus bradycardia | .279<br>(.007) | .280<br>(.008) |
| digoxin toxicity diagnosis and/or any symptoms | .417<br>(.007) | .399<br>(.009) |
| digoxin toxicity diagnosis | .003<br>(.001) | .003<br>(.001) |
| digoxin toxicity arrhythmia complex only | .075<br>(.005) | .075<br>(.004) |
| digoxin toxicity symptom complex only | .383<br>(.007) | .365<br>(.008) |
| Main Covariates | | |
| female | .736<br>(.007) | .735<br>(.008) |
| white*** | .667<br>(.007) | .732<br>(.008) |
| black*** | .070<br>(.004) | .030<br>(.003) |
| other race** | .262<br>(.007) | .238<br>(.007) |
| age 65-69*** | .262<br>(.006) | .236<br>(.007) |
| age 70-74 | .208<br>(.006) | .210<br>(.007) |
| age 75-79 | .148<br>(.005) | .160<br>(.006) |
| age 80-84 | .114<br>(.005) | .115<br>(.006) |
| age 85 or higher | .110<br>(.005) | .121<br>(.006) |
| length of baseline period*** | 10.94<br>(.037) | 10.74<br>(.048) |
| length of follow-up period*** | 13.27<br>(.093) | 12.69<br>(.114) |

Source: Washington Project C.A.R.E. Medicare/Medicaid Files
Note: *** statistically significant difference at the .01 level
** statistically significant difference at the .05 level
* statistically significant difference at the .10 level

**Exhibit 4.11.6**

**HCFA DUR Demonstration Evaluation**
**Summary of Signs of Treatment Effect Coefficients**
**Probit Models Based on Variables Created Using All Medicare Claims**

| Population | Dependent Variable | | | | | | |
|---|---|---|---|---|---|---|---|
| | Hyperkalemia | CHF | CHF+ | Digoxin Toxicity Symptoms or Diagnosis | Digoxin Toxicity Diagnosis | Arrythmia Complex | Symptom Complex |
| Main | -.033<br>(.075) | -.001<br>(.036) | -.025<br>(.034) | -.024<br>(.032) | -.051<br>(.158) | -.032<br>(.047) | -.030<br>(.033) |
| Baseline or Follow-Up Users | -.028<br>(.088) | -.005<br>(.044) | -.046<br>(.044) | -.026<br>(.042) | -.061<br>(.165) | -.067<br>(.056) | -.025<br>(.043) |
| Baseline Users only | -.003<br>(.097) | -.016<br>(.050) | -.049<br>(.048) | -.035<br>(.047) | -.009<br>(.179) | -.064<br>(.062) | -.017<br>(.017) |

Note: *** difference between arms statistically significant at the 1 percent level
** difference between arms statistically significant at the 5 percent level
* difference between arms statistically significant at the 10 percent level
*Sources:* Iowa Medicare claims

baseline congestive heart failure, baseline diuretics use, and baseline diabetes, as measured by use of insulin and oral hypoglacemics.

It was expected that the length of the baseline and follow-up periods might affect outcomes because individuals with longer baseline periods might have better information about baseline diseases and drug utilization, and because individuals with longer follow-up periods might have more time under observation. This hypothesis was confirmed in the cardiac effects and digoxin toxicity models; in the hyperkalemia models, however, baseline and follow-up length were not significant predictors.

The wave variables were created to capture learning effects and other variation in the provision and efficiency of cognitive services interventions that might be correlated with the pharmacy's date of entrance into the demonstration. In some models, it appeared that pharmacies that entered during the second wave had lower rates of outcomes compared to pharmacies that entered earlier or later. Generally, however, there appeared to be no consistent pattern between pharmacy wave and adverse outcomes related to cardiovascular drug use.

### Pharmacy Cluster Effects

To account for the fact that clusters of pharmacies rather than individuals were randomized to the treatment and control groups, each model included as a covariate the average rate of the dependent variable in the individual's pharmacy cluster. It was expected that this variable would help control for unobservable, cluster level differences in factors that affect the outcomes of interest. These variables did turn out to be statistically significant and positive in sign in most models. Nevertheless, the magnitude of the marginal effect of the cluster rates on the outcomes was usually very close to zero. Additionally, the Huber/White variance estimator, which also was used to adjust for cluster-based sampling, led to standard errors that were very similar to those calculated using a conventional variance estimate. Overall, then, adjustment for cluster-based sampling did not have any impact on the interpretation of the findings.

### Comment

Project C.A.R.E. was intended to encourage the provision of cognitive services by reimbursing pharmacists for these activities. The increase in cognitive services, in turn, was expected to reduce the rate of adverse clinical outcomes that result from inappropriate prescription drug use. This study focused on the impact of Project C.A.R.E. on a set of clinical outcomes related to cardiovascular drug use: (1) hyperkalemia; (2) congestive heart failure, hypotension, sinus bradycardia, and conduction disorders; and (3) digoxin toxicity diagnosis and clusters of symptoms related to digoxin toxicity. The analysis controlled for treatment and comparison group characteristics that may have differed at baseline, as well as the cluster-based sampling design used in the demonstration.

The findings indicate that Project C.A.R.E had no measurable impact on the outcomes of interest. It appears that the increase in cognitive services that resulted from reimbursement was not effective in preventing adverse effects related to inappropriate prescribing of cardiovascular drugs. There are a number of reasons why this effect might not exist, or why the effect might be too small to detect using this type of analysis. Project C.A.R.E. was not focused specifically on preventing adverse outcomes related to cardiovascular drug use; instead, the intervention was designed to target Medicaid users of any prescription drug. Furthermore, about half of cognitive services interventions that were performed focused on patient related problems, such as compliance. Only about 28 percent of cognitive services events resulted in a change in drug therapy [Christensen (1996)]. It is possible, then, that only a very small proportion of cognitive services events resulted in changes that even had the potential to prevent an adverse outcome related to cardiovascular drug use. This is particularly true for rare outcomes. Project C.A.R.E. therefore may not have effected measurable improvements in the rate of adverse effects related to antidepressant drugs.

It is more likely, however, that limitations in the study design clouded measurement of an existing treatment effect, particularly if the treatment effect was very small. Generally, a cluster sampling approach reduces the power of detecting differences between the treatment and control groups. In this case, however, the intracluster correlation is likely to be low; cluster based sampling therefore probably does not seriously reduce the power of the study design.

More important is the fact that the treatment effect is expected to be very small — it is hypothesized that the intervention might lead to half a percentage point difference in outcomes between the treatment and control groups. To capture such tiny differences, huge sample sizes would be needed. With the sample sizes used in the main population, the study design had less than 30 percent power to detect a half percentage point difference in outcomes in most cases.[66]

---

[66] The sample sizes required to detect differences of this size varied from 10,000 to 100,000 depending on the outcome.

## 4.12 Costs and Cost-Effectiveness of Project C.A.R.E.

### Measuring the Costs and Benefits of Project C.A.R.E.

The previous sections of this chapter have addressed the impact of Project C.A.R.E. on a number of adverse outcomes related to inappropriate prescription drug use. It was expected that Project C.A.R.E. would encourage pharmacists to offer cognitive services, and that the increase in cognitive services, in turn, would reduce the likelihood of potentially harmful drug use. Generally, the analysis did not support this conclusion; individuals who visited pharmacies where cognitive services were reimbursed were just as likely as individuals visiting other pharmacies to experience adverse outcomes associated with prescription drug use. Project C.A.R.E. therefore was found to be ineffective in improving prescription quality to the extent that adverse outcomes were averted.

Another aspect of the intervention, not discussed thus far in Chapter 4, is the economic cost of implementing Project C.A.R.E. If the evaluation had indicated that the intervention was effective, it would have been interesting to develop cost-effectiveness ratios and other measures that assess the benefits of the program relative to its costs. Nevertheless, despite the lack of evidence that Project C.AR.E. was effective, it is still interesting to consider the cost of operating the program. This information might be very useful to other states that are interested in initiating a cognitive services program. Moreover, even limited information on costs can be useful in gauging the magnitude of benefits that would be necessary to make the program an attractive option to policy makers.

The objective of this analysis was to use Project C.A.R.E. demonstration budgets to come to some understanding of the costs another state would incur if it were to implement such a program. As described below, the analysis was severely limited by a lack of separate budget information for research and operational tasks. For this reason, the costs of Project C.A.R.E. are discussed in general terms only and no attempt was made to calculate the average operations cost per cognitive service.[67]

### Data

The data used in this analysis come from Project C.A.R.E. demonstration budgets for fiscal years 1993-1996. The budgets were provided by the University of Washington. Each budget is divided into two sections: Washington and the University of Washington. The sections contain line items that were budgeted for each of these entities.

The University of Washington also provided expenditure information for cognitive services payments on a monthly basis. Otherwise, no expenditure information was available so the budgeted amounts were used as an estimate of what was spent.[68]

---

[67] For the Iowa OPDUR demonstration, some information was available that allowed a separation of operations and research costs in the demonstration budget. Chapter 5.13 presents an analysis that concludes with a presentation of the average cost of OPDUR per prescription screened.

[68] Dr. Dale Christensen, Principal Investigator, confirmed that this assumption was reasonable.

### Methodology

#### *Viewpoint of the Analysis*

The types of costs included in a study may vary by the perspective that explicitly or implicitly underlies the analysis. For example, if pharmacists spend large amounts of time learning how to provide cognitive services, this time is a cost from the perspective of the pharmacist, but it may not be a cost from the perspective of the government, the prescriber or the patient. In this analysis, a government payer viewpoint was adopted; costs were included only if they were borne by the government.

Ideally, the societal viewpoint would be most appropriate for this study because it is society as a whole that bears the costs and receives the benefits of Project C.A.RE.; the societal viewpoint implies that all costs, regardless of who incurs them, should be included in the analysis.[69] The societal viewpoint, however, suggests that a variety of cost data should be collected, including information from demonstration staff, patients, prescribers and pharmacists, and data on time costs and costs of donated materials. This analysis was limited to information available in the demonstration budgets. It was expected that other costs would be fairly small.[70] Nevertheless, because no information was available about these types of costs, the societal perspective could not be adopted in this analysis.

#### *Research and Operations Costs*

The demonstration budgets combine funds used for research purposes and funds used for Project C.A.R.E. operations. One of the central challenges of this analysis was to separate research costs from operations costs. Although Dr. Dale Christensen, principal investigator at the University of Washington, was very helpful in this task, it became clear that for many line items, it was not possible to calculate a reasonable estimate of what proportion of the line item was used for operational rather than research purposes. This problem was quite serious because without estimates of proportions of line items used for operations, the costs another state would incur for a cognitive services program could not be estimated. Instead, it is only possible to discuss generally the costs of important components of the program.

#### *Budgeted Versus Actual Expenditures*

The demonstration budget line items reflect funds that were granted by the Health Care Financing Administration. The funds may have been re-allocated between line items or not spent at all. This event may affect the analysis; in particular, if certain line items were not spent, the cost of Project

---

[69] It is true, however, that some groups bear more costs than others, and some groups enjoy Project C.A.R.E. benefits more directly than others. For example, a taxpayer who is not dually eligible for Medicare and Medicaid incurs Project C.A.R.E. costs, but receives benefits only indirectly, through the improved health outcomes of others.

[70] For example, the Project C.A.R.E. did not involve large quantities of donated goods or time, and it is unlikely that patients incurred any significant costs related to the demonstration. Pharmacists, and to a much lesser extent, prescribers, did spend time learning about Project C.A.R.E., but it is unlikely that this training took considerable amounts of time away from other work or leisure time.

C.A.R.E. might be over-estimated. Unfortunately, although expenditure information was available for cognitive services payments, no other information on expenditures was available.

### *Costs Specific to Washington*

Although the goal of this analysis was to estimate the cost another state would incur if they implemented a cognitive services program, some of the operating costs of Project C.A.R.E. are specific to the demonstration, and they would not be incurred by another state. As the principal investigator, Dr. Dale Christensen, pointed out, Washington uses a unique claims processing system rather than the universal claims system that most states use. Furthermore, at the time the demonstration was in operation, a national coding system for cognitive services could not be implemented, and the concept of a cognitive services intervention was still fairly new. If another state were to initiate a cognitive services program today, it would not be necessary to devote as many resources to training and the development of cognitive service codes and payment procedures as the demonstration did. Although it is difficult to estimate what proportion of the costs of Project C.A.R.E. could be averted by a state program implemented today, it is important to bear in mind that some of the costs of operating Project C.A.R.E. were unique to the demonstration.

### *Indirect Costs*

Project C.A.R.E. was primarily based at the University of Washington. The indirect costs of the demonstration therefore were calculated using rates set by the university. University indirect cost rates tend to be quite high, and they are not likely to be a good approximation of the indirect costs that would be incurred by a state implementing an cognitive services program. Furthermore, many of the indirect costs incurred in the demonstration may have been related to research rather than operations. It was not possible to estimate the operating cost portion of the indirect cost figure. For this reason, indirect costs are probably over-estimated in this analysis.

### *Time Frame*

The Iowa OPDUR demonstration project took place over three years, fiscal year 1993 - fiscal year 1996. Project C.A.R.E. was actually in operation, however, from February 1994 to September 1995. The time before and after the operations period was mainly allocated to research, start-up activities and phase-out activities, most of which would not be relevant to another state implementing a cognitive services program. For this reason, the analysis to a large extent was limited to costs incurred during the time period February 1994 to September 1995.

## Discussion

At the state level, Project C.A.R.E. costs included travel, data processing, payment for participation, payment for cognitive services provided, and staff time spent managing the demonstration. If another state were to implement a cognitive services program, it is unlikely that pharmacies would receive a participation fee. Furthermore, staff travel costs would be minimal if the program were operated by the state itself instead of by a university. Although some staff time would be require to answer pharmacists' questions and oversee program implementation, a state today would not face as many costs in this area as the demonstration did in 1994 and 1995.

The main, state-level Project C.A.R.E. costs that are relevant to another state, therefore, are the cognitive service fees and data processing costs. For the duration of Project C.A.R.E., approximately $52,234 was spent on cognitive services fees and $12,000 was spent on data processing (or $4.28 per CS and $5.84 per pharmacy). The amount of funds spent on cognitive services fees depends on the rate paid per intervention and the effectiveness of the program in increasing the provision of cognitive services. Both of these factors are likely to vary somewhat across states, at least in part because of different practice environments.

The cost of data processing also might vary across states. The process of integrating cognitive service payment codes into an existing claims database would require some programming and other staff costs in every state. The magnitude of these costs, however, would vary by the system the state was already using for claims processing. In the demonstration, the integration process was relatively expensive because of factors that are unique to Project C.A.R.E.; it is likely, then, that another state would incur lower processing costs than those incurred in the demonstration.

At the university level, the bulk of Project C.A.R.E. costs were allocated towards staffing needs, office equipment and space, and indirect costs. The majority of the staffing costs incurred by Project C.A.R.E. would not be relevant to a state implementing a cognitive services program. Many Project C.A.R.E. staff were involved in research and in developing and implementing cognitive services training. Today, it is unlikely that a state would offer such intensive training because pharmacists are much more familiar with cognitive services interventions now than they were at the time of the demonstration. Instead, a state simply might issue a memo describing cognitive services interventions. Furthermore, as mentioned earlier, the university indirect costs probably are high relative to a state's indirect costs. Overall, then, Project C.A.R.E. costs incurred at the university level are expected to be much higher than those that would be incurred by a state.

### Conclusion

The initiation of a state cognitive services program, similar to Project C.A.R.E., requires some start-up costs but it does not appear to be costly to maintain after implementation. The cognitive service payment system first must be developed conceptually and then integrated into the existing claims payment system. This process requires staff time mainly for management and programming, assuming that an existing system of payment codes is used. Additionally, a small amount of funds may be needed for training purposes.

After the system is in place, some staff time may be spent answering questions, revising codes as needed, helping to solve software problems, and resolving disputes about rejected claims. In many states, it might be possible to incorporate this staff time into an existing toll-free help line or other question/answer service. The cost of these services is likely to be quite low, but it was not possible to estimate their magnitude from Project C.A.R.E. budgets, which combined research and operational costs.

Although the limitations of this study precluded the calculation of a cost per cognitive service intervention, the analysis does reveal that there are significant costs of a cognitive service intervention that extend beyond the payments for the interventions themselves. This information is pertinent to states considering cognitive services programs, particularly because these costs might not be obvious in cases where tasks are integrated into existing systems. As the preceding discussion has indicated, to allocate resources in a cost-effective manner, it is worthwhile for states to consider the implementation and maintenance costs of a cognitive services program.

# 5.0 On-Line Prospective Drug Use Review: The OPDUR Demonstration

## 5.1 Background[1]

Medicaid is currently the nation's largest payer for retail prescription drugs. In the past few years, Medicaid and other third party payers have begun to take steps to protect patients from inappropriate prescribing and its potentially harmful consequences. One such step was the enactment of the Omnibus Budget Reconciliation Act of 1990 (OBRA 1990), which mandated that all state Medicaid programs implement prospective and retrospective drug use review (DUR).

Unlike retrospective drug use review, which focuses on reviewing the appropriateness of paid drug claims, prospective drug review begins at the point of sale, before a potentially inappropriate prescription is dispensed. Pharmacists initiate prospective DUR, reviewing prescriptions at the point of sale for potential problems and intervening with prescribers, other pharmacists or recipients as appropriate. Because complete information about patients' prescriptions is central to the success of prospective DUR, many states (23 in FFY 1995) use On-Line Prospective Drug Use Review (OPDUR), a computerized screening system that incorporates information from virtually all the pharmacies used by the patient.

Most pharmacies have computerized systems that screen prescriptions for potential problems. OPDUR is unique, however, in that submitted claims are screened against explicit criteria, and information on prescriptions filled at other pharmacies is available and incorporated into the screening process. The results of the screening procedure are transmitted back to the pharmacist within seconds; at the same time, eligibility verification and claims adjudication take place [Chrischilles (1997)]. If the pharmacist receives a message that the prescription did not meet the screening criteria (called a "screen failure" or "clinical event"), s/he is in a position to call the prescriber, meet with the patient and use his/her clinical judgement to resolve the problem.

Despite its widespread use, little is known about the effectiveness and cost-effectiveness of the OPDUR system. For this reason, the Health Care Financing Administration (HCFA) awarded Iowa a demonstration grant to study the OPDUR system. The demonstration award was granted in October, 1992, and Iowa began operating OPDUR in June, 1994. HCFA signed a agreement for the OPDUR demonstration with the Iowa Department of Human Resources, in cooperation with researchers from the University of Iowa, Drake University, and the Iowa Pharmacists Association. In turn, Iowa selected the National Data Corporation (NDC) to serve as the vendor of the OPDUR software. The Medicaid fiscal intermediary, Unisys, served as the host of the OPDUR software.

Although NDC provided the OPDUR software for the demonstration, Unisys and researchers at the University of Iowa spent considerable time and effort customizing the software, partly because of concerns about the product's specificity and positive predictive value. Extensive testing also was performed to ensure that the system was working properly, and to generate results from alternative parameter settings [Chrischilles (1997)].

---

[1] This section is based on the Iowa demonstration final report and reports to the TAG.

The main goals of the Iowa OPDUR Demonstration Project were the following [Chrischilles (1997)]:

- to create, test, and maintain an OPDUR system based on published criteria
- to assess the prevalence of potential drug-related problems detected by OPDUR
- to determine the impact of OPDUR on frequency and types of pharmacist interventions, drug use patterns, and Medicaid drug costs
- to explore patient, pharmacy, and OPDUR system characteristics that may modify the effect of an OPDUR system on cognitive services, drug use patterns, and drug costs

Pharmacies that wished to enroll were grouped into clusters based on shared prescribers. These clusters were then randomized to treatment and control groups. Pharmacies in the treatment group implemented the OPDUR system, and received OPDUR messages. Additionally, they were required to document cognitive services interventions. In the control group, pharmacies also implemented the OPDUR system and documented cognitive services interventions. The control group, however, did not receive OPDUR messages, even though their claims were screened by the OPDUR system ("passive screening"). This design allowed researchers to have equally detailed information for both the treatment and the control groups. The demonstration also included passive screening of claims from non-study pharmacies, to assess the possibility that merely participating in the study may have affected behavior [Chrischilles (1997)].

The Iowa OPDUR Demonstration Project eventually enrolled 223 pharmacies (110 intervention, 113 control) from 44 counties. Pharmacies were enrolled in two waves. OPDUR began operation for an initial set of 119 pharmacies in June, 1994; an additional 104 pharmacies were linked to the OPDUR system in December, 1994. The demonstration continued until June, 1997 [Chrischilles (1997)].

## 5.2 Iowa and Maryland Pharmacists and Prescribers: Attitudes Toward DUR and the Demonstration

In addition to analyses which measured the economic and health impact of the Iowa Medicaid OPDUR Demonstration Project, the Evaluator also engaged in process analyses to understand the context in which the demonstration operated. The process analysis drew on focus groups with pharmacists and prescribers (physicians and nurse practitioners), a survey of pharmacy owners and staff pharmacists, and targeted interviews during site visits. This chapter reports on findings from focus groups held in Iowa and Maryland in late 1995 and surveys conducted in 1996 in both states. Maryland was added to the evaluation to expand the effective size of the Iowa OPDUR Demonstration for analysis purposes. Maryland initiated its own statewide Medicaid OPDUR system in January 1993.

The focus groups were designed to better understand pharmacists' and prescribers' knowledge of and attitudes toward DUR systems, perceptions of their interactions with other health care professionals following OBRA 1990, and suggestions for improvements in prescribing. In Iowa, two focus groups were held, one for pharmacists and one for nurse practitioners and physicians. Four groups were held in Maryland as a result of more successful recruiting, two with pharmacists, one with physicians, and one with nurse practitioners. In Iowa, participating pharmacists were drawn from the general population, not from pharmacies participating in the Demonstration, so they did not address the Demonstration specifically. While focus groups do not produce findings that are generalizable either within or across states, participants' insights regarding DUR provide rich detail on aspects of the pharmacist role and pharmacist-prescriber-patient relationships regarding prescribing that can only be obtained in a group setting during open-ended discussion.

The survey collected information on pharmacy and pharmacist characteristics, pharmacist attitudes toward and experiences with DUR in the context of OBRA 1990 regulations, and, in Iowa, perceptions of the demonstration project. In Iowa, questionnaires were mailed to all pharmacy owners and pharmacists participating in the Demonstration; in Maryland, they were mailed to owners and staff pharmacists in about 200 pharmacies selected by a process similar to that which created the Iowa samples. Pharmacies with 50 or more Medicaid prescriptions per month were clustered around common prescribers; then clusters were selected for the sample at random until the target of 200 pharmacies was reached.[2]

This section of the report discusses Iowa survey and focus group responses to OBRA 1990 DUR requirements, attitudes toward and practices regarding OPDUR and the provision of cognitive services, implications for relationships between pharmacists, prescribers, and patients, and pharmacist attitudes toward providing DUR interventions in a changing health care environment. It does not report on all survey findings; a complete analysis of the Iowa and Maryland results can be found in the Evaluator's 1997 Annual Report.

Key findings were as follows:

- Focus group and survey findings strongly suggest that the provision of cognitive services is perceived by pharmacists and prescribers alike as important in ensuring proper prescribing and use of medication. For pharmacists, counseling is a burdensome responsibility that is difficult to conduct properly in pharmacies which demand ever-increasing volume;

- Prescribers asserted that pharmacy software that picks up drug duplication and interactions as the result of multiple pharmacy use and visits to multiple doctors would be of value to prescribers, who do not have that capability;

- Pharmacists and prescribers participating in focus groups in both states acknowledged that pharmacists are playing a greater role in the prescribing process. While pharmacists viewed themselves as partners and full members of the health care team in this role, prescribers were more ambivalent, acknowledging the value of the counseling role as a safeguard against prescribing errors and improper use of drugs but bristling when pharmacists overstep their bounds by recommending alternative therapies or advising patients against the use of certain drugs without consulting the physician first;

---

[2] Evaluation of Drug Use Review Demonstration Projects: Annual Report (1997). Cambridge, MA: Abt Associates, Inc.

- DUR was not perceived by pharmacists as the sole factor affecting prescribing behavior in recent years. Managed care, according to pharmacists in focus groups, was also perceived as a significant factor in the shift from pharmacist as drug vendor to pharmacist as drug expert. With increasing demands on their time, and given the wide range of drugs and formularies, physicians concede that they depend more on pharmacists to ensure prescribing safety. Consumerism is also cited by pharmacists as improving prescribing behavior;

- To improve prescribing, pharmacists recommended reimbursement for counseling services. Pharmacists and physicians advocated for linked computer systems that provide the pharmacist with a cross-pharmacy drug profile for the identification of abuse, therapeutic duplication, interactions, and other prescribing problems that the physician cannot obtain through patient history. Prescribers should have access to the same information.

### OBRA 1990 Requirements

OBRA 1990 required that state Medicaid programs subject all outpatient drugs to review. States were also mandated to implement retrospective DUR and patient counseling. Pharmacists across focus groups asserted that OBRA 1990 did not ask them to do anything beyond what should be expected of a pharmacist (e.g. counseling patients and contacting prescribers to clarify problems). Computerized DUR systems also existed prior to OBRA 1990, but to pharmacists, it was on-line claims *combined* with OPDUR that forced a change in pharmacy practice, according to Maryland pharmacists. Pharmacists spend more time on prescribing problems now in order to override claims rejections.

Pharmacists expressed several concerns about the OBRA 1990 requirements. Their liability should they fail to counsel patients has increased, so they have had to adjust their practices to ensure that they provide sufficient counseling. They have also had to purchase new software to implement OPDUR and electronic claims processing. These changes have added to their costs, so pharmacies have had to increase volume to make money, ironically leaving less time for counseling than in the past.

Physicians and nurse practitioners participating in focus groups were generally unaware of OBRA 1990 requirements, and in one group, nurse practitioners did not know what DUR stood for. One pharmacist recommended that prescribers be further educated on OBRA 1990 requirements and the changing role of the pharmacist in dispensing. He reported that he had visited the doctors with whom he most frequently interacted to explain the new regulations, but this was not common practice among his colleagues.

According to pharmacy owners responding to the survey, neither the operation of OPDUR nor the provision of counseling to Medicaid recipients and caregivers represented an unusually large financial burden on participating pharmacies. Despite the low *percentage* of costs attributed to ProDUR and counseling, a majority of Iowa pharmacy owners indicated an increase in prescription costs over the previous year attributable to ProDUR and counseling, while Maryland pharmacy owners saw their prescription costs increase less often than in Iowa. It may be that, to a great extent, pharmacists are absorbing the additional demands on their time as indicated by focus group participants who described themselves as dedicated to counseling despite working in a higher volume, less personalized environment. One Maryland pharmacist, though, described the tension between achieving dispensing volume goals and providing quality counseling and concluded that, in a chain pharmacy environment, volume takes precedence over counseling.

Pharmacists, nurse practitioners, and physicians participating in focus groups discussed the expanded role of the pharmacist in the prescribing process as a result of OBRA 1990. Pharmacists responding to the survey generally saw their role as an important one. Most pharmacists agreed or strongly agreed that without monitoring patient drug therapy, an unfavorable outcome is likely (81 percent). Almost all pharmacists (94 percent) agreed or strongly agreed that patient care would suffer if they failed to provide drug information to physicians.

Physicians in both Iowa and Maryland focus groups observed that the pharmacist is becoming more proactive as indicated by the increased frequency of their phone calls on a variety of prescribing issues. Pharmacists felt that OBRA 1990 had made them feel more like health care providers than dispensers of medication, made them more accessible to the customer, and made them more concerned about liability:

> "...if OBRA '90 was not there, I couldn't remember to have to consult half these patients. My techs give it out, they don't know what's going on."
> "OBRA's forced me to spend more time with the customer."
> "Maybe we're now dispensers of knowledge, probably, instead of the medications themselves."
> "We're like that last link to the patient, and I think the patients have come to count on the pharmacist a lot more for information." (Maryland pharmacists)

Pharmacists also see physicians as more dependent on them in the last few years, but they attribute this shift to managed care, the huge variety of prescription drugs on the market, and the variability of formularies by insurance type. Managed care was viewed, particularly in Maryland, as reducing the amount of time doctors had with their patients, "so it's going to mean more responsibility for us, whether it's due to OBRA or not" (MD pharmacist). Physicians, one pharmacist suggested, may also appreciate the pharmacist picking up on errors more than they used to due to liability concerns.

OBRA 1990 has made some pharmacists examine their role in relationship to that of the pharmacy technician. Some pharmacists use the technician for filling the prescription while they do the computer profiling and counseling. This arrangement ensures that the pharmacist is accessible to patients at all times and ensures that counseling takes place. Other pharmacists have the technician at the computer, alerting the pharmacist to problems; the pharmacist then counsels the patient. It appeared to be the consensus of pharmacists that it was essential for them to do the counseling themselves.

One finding from the Iowa OPDUR Demonstration underscored the importance of counseling relative to the use of the computer to detect prescribing problems. Although the primary focus of the Iowa Demonstration was OPDUR, researchers required participating pharmacists to document cognitive services as well. Only 31 percent of the prescribing problems cited on the CS forms were those picked up by the computer [Chrischilles et al. (1997)]. Forty percent of the pharmacists' interventions would have gone undetected, the authors state, if they were solely dependent on the computer. This finding, along with the importance placed on counseling by pharmacists in the focus groups and survey, implies that while OPDUR detects many (not all) prescribing problems, the counseling process both identifies problems and intervenes to prevent misuse or abuse of prescription drugs. It also emphasizes the need for a trained pharmacist to conduct the counseling.

### Attitudes Toward OPDUR

The survey attempted to measure pharmacists' attitudes toward OPDUR. It was hypothesized that Iowa-treatment pharmacists would report a more positive attitude toward OPDUR than would Maryland pharmacists because the Maryland system had been imposed on them, but the survey found no significant differences across groups. The Iowa-treatment group, however, scored slightly higher on overall attitudes toward OPDUR than Maryland pharmacists.

Most pharmacists in Maryland and in the Iowa-treatment group, both of which had OPDUR systems in place, reported a positive attitude toward the OBRA 1990 requirements for the implementation of ProDUR. They reported that OPDUR does not cause problems in their relationships with other health care professionals. Most agreed or strongly agreed that OPDUR did not interfere with the patient-prescriber relationship (71 percent), the pharmacist-prescriber relationship (74 percent), or the patient-pharmacist relationship (83 percent). Most pharmacists agreed (77 percent) that OPDUR helps avoid serious adverse effects.

It was hypothesized that pharmacist workload (as measured by prescriptions filled in an eight-hour shift) would predict attitudes toward OPDUR and cognitive services. Maryland pharmacists reported a much higher workload than did Iowa pharmacists, yet their attitudes toward OPDUR, as reported in the survey, were generally positive. When queried about the benefits of OPDUR, focus group participants suggested that the OPDUR system eased their load by providing them with the backup they needed to ensure that they were identifying prescribing problems during a hectic day. The computer, pharmacists noted, was not only useful for the identification of prescribing problems but for patient profiling.

Iowa pharmacists-in-charge were asked questions related to the OPDUR Demonstration. Most respondents showed general satisfaction with communications they received about cognitive services documentation and about OPDUR. They recommended that outreach and reminder phone calls, newsletters, and in-person visits from project staff would motivate them to document cognitive services.

### Attitudes Toward Cognitive Services

On average, pharmacists completing the survey reported higher levels of satisfaction with the provision of cognitive services[3] than with OPDUR. Overall, 88 percent of pharmacists agreed or strongly agreed that the provision of cognitive services helped avoid serious adverse patient effects. Iowa-treatment pharmacies overall felt the most positive about the provision of cognitive services whereas Maryland pharmacists were the least positive (using a combined score on a series of attitudinal questions). Iowa pharmacists responded that they were more likely to be able to provide as much counseling as they felt was needed than were Maryland pharmacists, which may explain their more positive attitude.

For pharmacists who were not able to do as much counseling as they would like, several reasons were cited in open-ended survey responses:

- pharmacists spend so much time dispensing they have little time to counsel.
- concern about the financial feasibility of counseling. As stated by one pharmacist: "lower levels of reimbursement by third parties put pressure [on me] to fill more prescriptions with [the] time available."
- the person picking up the prescription is not the beneficiary.

All of these findings were supported by the focus groups, in which pharmacists spoke at length about barriers to providing counseling, from financial concerns to higher workload, to patient-specific factors such as language differences and rejection of the offer of counseling.

A series of questions in the survey of pharmacists were aimed at determining whether pharmacy characteristics such as volume, type of pharmacy (independent vs. chain), geographic location, or the existence of a separate counseling space affected pharmacists' attitudes toward cognitive services. It was found that only volume was associated significantly with attitudes toward cognitive services; as volume increased, attitudes toward CS were more negative, which is consistent with the focus group findings.

As for cognitive service documentation (only required in Iowa), as workload increased, the burden of documentation was seen as greater. Overall, a negative relationship was found between workload and attitudes toward cognitive services which was of borderline statistical significance; as workload increased, there was a slightly more negative attitude toward cognitive services. Focus group participants stated that they were not required by any insurer to document cognitive service provision so were unable to discuss this component of the counseling process.

### Payment for Provision of Cognitive Services

One of the key areas of interest as DUR programs are implemented is whether or not paying pharmacists for documenting the provision of cognitive services would act as an incentive for them to provide consistent, standardized counseling services to their clients with Medicaid. At present, pharmacists in Iowa and Maryland are not paid to provide cognitive services to Medicaid beneficiaries. Only four percent of pharmacists in Maryland and Iowa responding to the survey reported being paid from any source for the documentation of cognitive services.

---

[3] The survey used the term "cognitive services" as it was defined in the OBRA 1990 legislation to describe the following pharmacist activities: instances when a pharmacist consults a prescriber (or other pharmacist); patient or caregiver counseling; monitoring or educating patients; consulting external sources of prescription drug information. Focus groups referred to these activities more generically as "counseling."

It is clear from the focus groups that pharmacists are struggling with increased costs and workload, yet because they see counseling as integral to ensuring that patients take the proper medicines in an appropriate fashion, they provide the service. However, they expressed concern that counseling could easily fall by the wayside in such a work environment. They fear that computer printouts, known to some as "written counseling," are being substituted for in-person counseling. These documents, handed to patients along with their prescriptions, describe the drug, its administration, and possible side effects. While informative to consumers, printouts are perceived as controversial. They list *all*, not just the most common side effects, thereby raising the anxiety level of patients who return to their doctors, hesitant to take the drug. Physicians agreed that the printouts sometimes cause unnecessary anxiety among their patients and damage their credibility with patients who feel as if they were not fully informed of the consequences of using the drug.

The printouts can precipitate tension between pharmacists and physicians. As one Maryland pharmacist noted that physicians complained that their patients were frightened by the printouts. On the other hand, pharmacists feel that the printouts are useful in that they fulfill the mandated task of counseling and give consumers more information. They do not, however, substitute for the more valuable face-to-face counseling that is perceived as much more comprehensive and effective, particularly with low-income patients, persons on multiple medications and those whose primary language is not English. Pharmacists in the focus groups argued that they should be compensated for the time they spend counseling and contacting prescribers, particularly as prescribers increasingly depend on them to be the drug experts and safety net for patients on complicated drug regimens.

### Prospective DUR and the Pharmacist-Prescriber Relationship

It is apparent from the qualitative data that, to pharmacists, physicians, and nurse practitioners, the role of the pharmacist is expanding. While this change is viewed positively by pharmacists, physicians and nurse practitioners expressed mixed feelings. Prescribers viewed pharmacists as most helpful in identifying duplicate drug prescribing and drug-drug interactions. With the capacity to obtain drug information from multiple pharmacies, pharmacists can determine whether or not the patient is receiving two similar prescriptions or drugs that interact. Often, the prescriber is unaware that the patient is seeing multiple health care providers and has no knowledge, other than that reported by the patient, of other drugs that are being prescribed.

Some physicians in focus groups reported that pharmacists have become more aggressive and adversarial, challenging the physician's authority. They expressed concern that pharmacists, without consulting them first, caution the patient against using the drug when the physician has expressly, and with knowledge that this is not the conventional indication for the drug, prescribed it. Physicians see this as interference, unless the pharmacist contacts them first to discuss the issue at hand. Other physicians commented that pharmacists convey their prescribing concerns well and at times act as team members on behalf of the patient.

The changing role of the patient from passive recipient of health care to consumer as well as managed care have worked simultaneously to change physician behavior, according to physicians in Maryland focus groups. These dual trends may, they acknowledged, lead to better prescribing. "Patients are beginning to expect more, to demand more," thereby driving physicians to depend on pharmacists as a backup to ensure proper prescribing. Both in response to consumer demand and pharmacists, physicians in Maryland asserted that they are being forced to give patients more drug information before leaving the office so "we won't look like fools." Managed care, pharmacists say, gives physicians less control over what they prescribe, so they are seeing prescriptions for smaller quantities and less expensive therapeutic choices.

In terms of how pharmacists view physicians, pharmacists in Maryland view themselves as more accessible to patients than doctors and therefore hear more of patients' concerns and fears regarding prescribed drugs. They also hear more about lapses in taking medications which they can then convey to the physician. Pharmacists in this group did not believe that physicians always prescribe the best and most cost effective medications, in part because they are unaware of less costly options and do not know formularies.

### Recommendations

Pharmacists across focus groups strongly recommended that they be paid for counseling. Some pharmacists recommended increased reimbursement for counseling to allow them to provide this service exclusively, leaving dispensing to technicians. In fact, it was suggested by one pharmacist that counseling of certain patients such as diabetics should extend beyond drug regimen to proper diet and use of syringes. Pharmacists acknowledged that in order to survive, they cannot depend on dispensing fees alone. They urge the execution of outcome studies which investigate health outcomes of enhanced counseling. Physicians recommended that inter-pharmacy computer linkages be expanded and that they, too, obtain access to all DUR information on-line.

### Conclusion

Pharmacists conveyed positive attitudes in both the survey and focus groups toward both OPDUR and counseling. They use the two in combination to identify possible prescribing problems and intervene with patients and prescribers to avert therapeutic errors, misuse, and abuse of prescription drugs. Physicians and nurse practitioners, overall, expressed positive feelings about the use of computerized DUR, particularly its ability to detect duplicate prescriptions and interactions that result from multi-pharmacy use and multiple providers. The expanded role of the pharmacist into the counseling realm is perceived as positive except when the pharmacist questions the safety or indications for a particular medication with the patient without consulting the prescriber first. To pharmacists, the optimal DUR system appears to consist of OPDUR with multi-pharmacy capability combined with electronic claims processing, and reimbursable counseling.

## 5.3 The Iowa OPDUR Demonstration: Implementation

The implementation of the Iowa OPDUR demonstration involved a number of steps, beginning with the recruitment and enrollment of pharmacy participants, and ending with a continued effort to retain the sample and maintain the OPDUR system. These steps are described in this section. As in the

Washington demonstration, implementation of the Iowa OPDUR demonstration brought with it various challenges, some of which required a modification of original plans. In general, however, the implementation of the demonstration went smoothly.

### Recruitment, Enrollment, and Maintenance of Pharmacy Participants

Iowa pharmacies were eligible for participation in the OPDUR demonstration if they met at least two of the following three conditions: (1) had at least 1000 non-nursing home pharmacy claims in 1992; (2) had at least 1300 pharmacy claims in 1993; and (3) had at least 1000 projected point-of-service (POS) claims for 1994. Pharmacies without complete information were evaluated using 1993 claims information or projected 1994 information. Pharmacies with relatively low claims volume were excluded because the expected rate of interventions was quite low (1-2%) [Chrischilles (1997)].

Pharmacies were encouraged to participate in the demonstration through newsletters, recruitment letters, and articles in the Iowa Pharmacists Association Journal. Pharmacists who indicated that they might be interested in enrollment were invited to training sessions. At the training sessions, the goals and objectives of the demonstration were described, and informed consent was obtained [Chrischilles (1997)].

Efforts were made to train new (second wave) entrants to the study as well as to maintain the first wave participants. The demonstration offered additional training to new pharmacists at participating pharmacies, and participating pharmacies were sent outreach newsletters. A help desk was maintained so that pharmacists could call and ask specific questions about OPDUR messages and screen failures. Participants with more general questions were invited to call the University of Iowa research office.

### OPDUR Implementation and Maintenance

To implement the OPDUR screens, Iowa employed an "integration product" that adapted commercial software to local conditions. NDC tailored existing, commercially available on-line prospective DUR software, known as the ProDUR™ system, to meet the objectives of the Iowa demonstration. While the ProDUR™ software included a set of clinical modules that were composed of common problems (e.g. drug interactions, duplicate drug therapy, therapeutic overlap, age precautions etc.), the demonstration project determined the threshold parameters and the specific drugs to be screened. In addition to three screening criteria specified in the UM/PCPS criteria (dosage, drug-drug therapeutic duplication, and drug-drug interaction), Iowa also screened for overuse/early refill of all drugs. Local modifications included the incorporation of a drug use chronology, modifying disease precaution and duration conflict screens if the dosage is above/below a certain threshold, and adding age modifiers when appropriate [Chrischilles (1997)].

The demonstration project and Unisys took on much of the responsibility for customizing ProDUR™. This responsibility included tasks such as defining the syntax for modifications, and developing software to operationalize the syntax [Chrischilles (1997)]. Before the OPDUR system could be linked to pharmacies, several screening problems had to be resolved. These problems ranged from denials not being screened to problems in generating reports of screen failures [Chrischilles (1997)].

Pharmacies in the first enrollment wave were linked to an operating OPDUR system on June 13, 1994. Initially, demonstration project staff called pharmacies who should have been receiving OPDUR messages to ensure that the messages had been transmitted properly. At first, a large percentage of participants (40%) experienced problems in receiving messages. By August, 1995, however, all but one intervention pharmacies were receiving messages properly.

The OPDUR system required regular maintenance. Every two weeks, the demonstration received updates to the ProDUR™ system which were evaluated, and, when deemed relevant, incorporated into the OPDUR system. System performance was checked on a monthly basis and screen failure data was downloaded weekly. Throughout the demonstration, NDC, Unisys and the University of Iowa worked together to resolve quality control and system modification issues [Chrischilles (1997)].

### Iowa OPDUR Demonstration Findings

The Iowa OPDUR Demonstration investigators [Chischilles, (1997)] reported little direct evidence that OPDUR affected pharmacist behavior, though analyses suggested that there was a slightly higher rate of claims reversals in treatment than in control pharmacies.

- In the Iowa demonstration, CS rates declined over time . Low volume pharmacies delivered more CS than high volume pharmacies. Mean CS time was 9.4 minutes, with control pharmacies spending more time (12.1) than treatment pharmacies (6.4). The direction of difference remained the same and was statistically significant when controlling for problem, intervention and result type.

- The average rate of documented CS per 100 claims in the Iowa OPDUR demonstration was 0.19 (0.16 for treatments, 0.21 for controls) CS per pharmacy was higher for control pharmacies (36) than treatment pharmacies (27.8). An important difference appeared to be that control pharmacists recorded CS related to non-formulary/prior authorization issues more frequently than did treatment pharmacists. In general, controls were more likely to document CS not related to the OPDUR screens. Iowa investigators suggested that part of the difference in reporting rates might be due to "compensatory rivalry," with the controls attempting to show that they were as capable or more capable of delivering CS. Investigators also speculated that the difference might have been due to the different "triggers" presented to the two groups (screen failures for treatments, prior authorization/claims rejections for controls). But 40 percent of all CS was unrelated to Medicaid formulary, prior authorization or computer-assisted DUR.

- In Iowa, 47.6 percent of documented CS were provided for "patient" issues. Control pharmacists were more likely to consult a prescriber (82 percent) than treatment pharmacists (64 percent). The reverse was true for CS that involved consulting a patient.

  Therapy change occurred in 55.1 percent of CS, and the rate was higher for control than treatment pharmacists (60.1 percent, compared to 48.3 percent.). The odds of CS leading to therapy change were 1.3 percent higher for control pharmacists, adjusting for documented problem and interventions.

- In the Iowa OPDUR Demonstration, 7.4 percent of all claims generated at least one screen failure. Early refills were the most frequent failures (4.1 percent of all claims), whereas drug-drug interactions were detected for only 1.2 percent of all claims. Among elderly beneficiaries, over one-third had at least one screen failure per quarter, with early refills the most common problem.

- Claims reversals can be thought of as a partial indicator of drug therapy change, less affected by possible reporting biases than CS. Iowa investigators found that the percent of claims reversals associated with at least on screen failure increased over the demonstration, and was somewhat higher on average for treatment (4.2 percent) than control pharmacies (3.7 percent). Differences when more than one pharmacy was involved were of similar size, in the hypothesized direction (a higher rate for treatment pharmacies) and statistically significant. For multiple prescribers, however, though the difference was in the right direction, it was not statistically significant.

## 5.4 Findings on the Impact of the OPDUR Demonstration on Downstream Utilization and Costs

### 5.4.1 Background

This section of the report evaluates the Iowa On-Line Prospective Drug Use Review (OPDUR) demonstration, which took place between June 1994 and June 1997. In theory, the availability of a system that instantly screens prescriptions for possible errors should improve pharmacists' knowledge base and increase their efforts to ensure optimal drug utilization engaging in more "cognitive services" (CS) activity, improving drug prescribing and patient compliance, and thereby decreasing Medicaid recipients' "downstream" health services utilization (specifically utilization involving diagnoses related to suboptimal drug use). Section 4.6 reported essentially null findings from analyses of downstream utilization effects in Washington's Project C.A.R.E. demonstration of (CS) payment.

The evaluation team tested the hypothesis that access to an OPDUR system caused a decrease in non-drug health services utilization among individuals aged 65 and over who were dually eligible for both the Medicare and Medicaid program. This hypothesis was also tested for dually eligible individuals under 65, and for subpopulations of the elderly thought to be at risk for suboptimal utilization (diabetics, individuals with three or more prescribers, individuals using ten or more prescription drugs). Particular attention was paid to downstream effects on targeted drug/utilization categories, defined by the drug group criteria included in Iowa's OPDUR system. As in the Project C.A.R.E. analyses reported above, the Iowa analyses found no evidence to support the downstream utilization hypothesis.

### 5.4.2 Data and Methods

This study investigated whether Medicaid recipients who filled their prescriptions at treatment pharmacies had lower health services utilization than recipients who filled their prescriptions at control pharmacies. For such an effect to exist, two things must have happened. First, access to OPDUR must have led pharmacists to perform more cognitive services; even without such additional incentives, pharmacists make efforts to guarantee that Medicaid recipients receive safe and effective drug therapy.

Second, these additional services must have measurably affected utilization. Either they must have substituted directly for other utilization or they must have been additional services with a protective effect. Either by changing prescriptions or by changing recipients' behavior, cognitive services must have improved subsequent health and reduced subsequent health services utilization.

The purpose of this study was not only to evaluate whether there was a general effect but also to find and explore any specific effects that access to OPDUR might have had. Thus, the study emphasizes aspects of utilization, defined by diagnosis, and populations expected to be especially sensitive to such payment.

### *Selection of "At Risk" Populations*

The base population for this study was individuals age 65 and over, eligible for both Medicaid and Medicare, not enrolled in managed care and not resident in nursing homes, who filled prescriptions at demonstration pharmacies. The elderly were chosen primarily because the they are, on average, in relatively delicate health and receiving relatively many prescription drugs. Thus, they are at greater risk for drug-related health problems and were hypothesized to be more likely to benefit from cognitive services than younger populations. Dual eligibles were chosen because the intervention was directed at Medicaid recipients, the vast majority of elderly Medicaid recipients are also enrolled in Medicare, and the data would not be consistent if dual eligibles were mixed with those who were only eligible for Medicaid. In addition, four other subpopulations were defined for the analysis:

- Dually eligible individuals under age 65;
- Dually eligible elderly with diabetes;
- Dually eligible elderly with three or more prescribers:
- Dually eligible elderly with ten or more prescriptions

For the analyses of utilization in diagnoses known to be related to certain prescribing problems, the diagnoses were aggregated to the level of the drug class (for example, all diagnoses know to be related to less-than-optimal prescribing of NSAIDs) and the study population was subset to individuals who used drugs in that class. Drug classes were those included in the PSU screener criteria.

Finally, it can be argued that downstream utilization effects are most likely to be concentrated in the subset of treatment pharmacy users for whom OPDUR screens have been triggered and messages generated. Utilization of these individuals is compared to utilization of the at risk subpopulation of control pharmacy users; i.e. , those whose prescriptions triggered OPDUR screens without generating corresponding messages to their pharmacists.

### *Data Base and File Structure*

The data for this study were the Medicaid drug claims and the Medicaid and Medicare health services utilization claims for individuals in the state of Iowa in 1993 (to provide a baseline) and 1994-97. (The demonstration took place between June 1994 and June 1997.)

Individualized baseline and follow-up periods were constructed for each sample member. Follow-up periods began on the day that the person first went to an active demonstration pharmacy and lasted until June 1997, loss of eligibility for either Medicaid or Medicare, entry into a nursing home, or death, whichever came first. Similarly, baseline periods ended the day before the person first went to an active demonstration pharmacy and stretched back 12 months or to the first month the person was eligible for both programs or not in a nursing home, whichever came last. These windows were intended to capture the period when the demonstration could and could not have affected health services utilization and when Medicaid and Medicare claims would offer a complete record of health services utilization. The study population included only individuals with positive follow-up time, this being the population that would exhibit intervention effects, and at least three months of baseline because many important covariates were measured in the baseline period. (Results from analyses of those individuals lacking baseline data were broadly similar to those reported here.) The final samples also excluded the relatively few individuals who filled prescriptions in both treatment and control pharmacies.

The elderly population was further subset to groups especially likely to benefit from the demonstration. Diabetics were identified via claims for insulin or oral hypoglycemics at any point in 1993-1997; individuals with multiple prescribers via claims from at least three physicians in a single baseline month; individuals with complex drug regimens via claims for at least 10 distinct drugs, as indicated by distinct NDC-9 codes, in a single baseline month. In addition, the population of dually-eligible non-elderly persons (adult disabled) was included in the analyses.

For some analyses, the population was subset to users of a particular drug, based on use during the baseline period This guaranteed that the defined subpopulations were independent of intervention effects, but excluded individuals who began using screener drug during the demonstration. This could be a serious omission if, as seems likely, new users are particularly good candidates for adverse drug reactions. Therefore analyses were repeated using an alternative sampling frame that induced individuals who used drugs in the class during either the baseline or the follow-up period. Results for this second population were similar to those for the first and were omitted for brevity.

As Exhibit 5.4.1 shows, beginning with all Medicaid eligible persons who filled prescriptions at any demonstration pharmacy between 1993 and 1997 (145,764), analysts eliminated those with inadequate baseline or follow-up data (66,412), those who could not have been exposed to the demonstration (959) and those who became 65 during the study (818). Individuals who filled prescriptions in both treatment and control pharmacies (7,951) and individuals with inadequate demographic data (4,376) were also eliminated, leaving a "net usable study population" of 65,248. Of this group, 14 percent (9,075) were dually eligible with at least three months of baseline data, and from among these individuals, five analysis samples were created: a) all elderly (4,597), b) elderly diabetics (951), c) elderly with three or more prescribers (990), d) elderly with ten or more prescriptions (489) and e) non-elderly (4,478).

### *Utilization Measures*

Downstream utilization measures included indicators both of whether or not a service was utilized and, for users, how much, usually in terms of total combined Medicare and Medicaid payments. Exhibit 5.4.2 describes these measures. They include inpatient and professional services, as well as

**Exhibit 5.4.1**

**HCFA DUR Demonstration Evaluation**
**Derivation of Iowa OPDUR Demonstration Downstream Analysis Samples**

| | | |
|---|---|---|
| 1. | Medicaid eligible, filled prescription at demonstration pharmacy, 1993-1997 | 45,764 |
| | *Less* | |
| | Individuals with insufficient data (no follow-up period, baseline less than three months) | (66,412) |
| | Not exposed (timing of first prescription) | (959) |
| | Became 65 during study | (818) |
| | Filled prescriptions in both arms | (7,951) |
| | Inadequate demographic data | (4,376) |
| 2. | Net usable study population | 65,248 |
| | Dually eligible (Medicaid/Medicare) | 9,075 |
| | **Elderly (65+ at beginning of study, 9/30/95)** | **4,597** |
| | **Elderly diabetic** | **951** |
| | **Elderly with 3+ prescribers** | **990** |
| | **Elderly with 10+ prescriptions** | **489** |
| | **Non-elderly (<65 at end of study, 6/30/97)** | **4,478** |

Source: Iowa OPDUR Demonstration Medicare/Medicaid Files

**Exhibit 5.4.2**

**HCFA DUR Demonstration Evaluation**
**Utilization Measures: Iowa OPDUR Demonstration Downstream Analysis**

| Measure | Description |
|---|---|
| Inpatient (0,1) | Probability of an inpatient admission reimbursed by Medicare or Medicaid |
| Inpatient ($) | Total Medicare and Medicaid payment, including professional services, for hospital admissions |
| Professional (0,1) | Probability of a professional service encounter (single day's interaction with a professional provider) delivered in a non-hospital setting |
| Professional ($) | Total Medicare and Medicaid payment for professional service encounters |
| ER (0,1) | Probability of an outpatient emergency room encounter |
| ER($) | Total Medicare and Medicaid payment for outpatient emergency room encounters |
| Total (0,1) | Probability of any Medicare or Medicaid service* |
| Total ($) | Total Medicare and Medicaid payment for all services* |

* Includes inpatient, professional, emergency room, and skilled nursing facility services

Source: Iowa OPDUR Demonstration Medicare/Medicaid Files.

outpatient emergency room admissions. ER measures were expected to capture acute prescription-drug related episodes treated in an outpatient rather than an inpatient setting. [4] The study employed both broadly-defined measures (for example, the frequency of all inpatient admissions in the demonstration sample) and more narrowly -defined measures ( the frequency of inpatient admissions for diagnoses linked to less than optimal prescribing of specific drugs). To support these more targeted measures, researchers at Penn State University assembled a data base of known relationships between clinical outcomes and selected prescribing problems, namely those problems that are included in the University of Maryland/Philadelphia College of Pharmacy Science screener. This screener covers four prescribing problems — dosage, drug/drug interactions, duration of therapy, and therapeutic duplication — and eight classes of drugs — ACE inhibitors, anti-depressants, anti-psychotics, benzodiazepines, digoxin, calcium channel blockers, H2RAs, and NSAIDS.

The PSU outcomes data base is comprehensive and well-documented. It is the result of a search of primary literature via electronic bibliographic data bases, a manual search of selected secondary sources such as the major drug compendia and resource texts, and extensive review and comment from clinical experts. Within that data base, targeting diagnoses were selected for this study if both the diagnosis and the associated prescribing problem were sufficiently prevalent in the Iowa Medicaid population that the study would have sufficient power to detect a demonstration effect of moderate size (in the neighborhood of one quarter of one standard deviation).

### Independent Variables

In modeling the effects of the demonstration on downstream utilization, analysts first developed measures to characterize demonstration exposure and then selected, from the available data, other covariates that theory suggests may have utilization impacts separate from those due to exposure. The independent variables selected for this model are listed in Exhibit 5.4.3.

#### *Characterization of Demonstration Exposure*

Although this demonstration was designed as a random experiment, randomization occurred at the level of the pharmacy while many of the benefits of payment for cognitive services were expected to accrue at the level of the pharmacy patron/Medicaid recipient. This posed several modeling challenges. First, individuals had the potential to go to multiple pharmacies and, potentially, to go to a mix of intervention, control, and non-demonstration pharmacies. The study adopted a simple approach. Individuals were considered members of the intervention group if they ever went to an intervention pharmacy and never went to a control pharmacy. They were considered members of the control group if they ever went to a control pharmacy and never went to an intervention pharmacy. Individuals who went to both intervention and control pharmacies were excluded from the study.

---

[4] Analysts studied skilled nursing facility utilization, and measures of total utilization and expenditure include SNF services. However, SNF admissions in all of the analysis populations were infrequent (1.6 percent among all dually eligible elderly sample members), and absent altogether for some of the smaller samples. SNF utilization was, therefore, not included as a separate dependent variable.

**Exhibit 5.4.3**

**HCFA DUR Demonstration Evaluation Independent Variables: Iowa OPDUR Demonstration Downstream Analysis Samples**

| Variable | Definition |
|---|---|
| *Demonstration exposure* | |
| Treatment (0,1) | Demonstration status of individual: *treatment* if all prescriptions during study period filled at treatment pharmacies, *control* if all filled at control pharmacies |
| *Other independent variables* | |
| Age | Nine qualitative variables: 20-29, 30-39, 40-49, 50-59, 60-64, 65-69, 70-79, 80-84, 85+ |
| Sex | |
| Race | Three qualitative variables: white, black, other |
| Eligibility | Medicaid eligibility - blind or disabled |
| Baseline utilization | Value of the dependent variable during baseline period |
| DCG/HCC | Predicted annual Medicare payments |
| Baseline drugs | Maximum number of NDC-9 level distinct drugs in a single month during baseline period |
| Length of baseline period (months) | |
| Length of follow-up period (months) | |
| Wave | Recruitment wave of pharmacy filling individual's first prescription (three waves in Washington) |

Source: Iowa OPDUR Demonstration Medicare/Medicaid Files

In order to calculate and apply the cluster-level averages, it was necessary to assign individuals to clusters. This assignment was based on the cluster affiliation of the first active demonstration pharmacy to which the person went.

The intervention/control indicator was admittedly a simplification of a potentially complex exposure. Individuals deemed members of the intervention group varied in terms of the number of prescriptions they filled at intervention pharmacies, the complexity of their drug regimens, the length of time they went to intervention pharmacies, and whether their exposure was relatively late in the demonstration. One might expect the intervention's impact to be increasing in all of these dimensions. Members of the intervention group also differed in age, and a hypothesis of the evaluation was that demonstration effects would be especially pronounced among the elderly. To capture this complexity, the study experimented with alternative specifications featuring interactions of intervention status with appropriate variables (always also incorporating both the intervention status indicator and the variable itself). No consistent significant results emerged, and results of these interactions aren't shown.

A second modeling issue was that the regression framework required adjustment in order to account for this experimental structure. The number of pharmacy clusters was fairly small leading to differences in measured demographics and baseline outcome rates between intervention and control recipients. While one could control for any measured differences directly within the regression framework, the presence of these differences raised the suspicion that there could also be unmeasured differences in demographics, access, or medical practice across clusters that would be correlated with intervention status and also with some or all of the dependent variables. Because the number of clusters was small relative to the number of recipients, these differences might not average out between intervention and control groups despite the large number of recipients leading to misleading estimates of intervention effects unless an appropriate adjustment was made. Specifically, traditionally-estimated standard errors would be smaller than the true standard errors.

The first adjustment was to include the cluster-level averages and demographics discussed above. These variables controlled directly for measurable differences among individuals and actually improved the precision of estimated intervention effects. The second was to employ the econometric technique known as variable effects; this technique effectively adjusted the standard errors of estimated parameters to account for the number of clusters and the amount of the marginal variation in the dependent variable that they explained.

### *Other Covariates*

The Andersen/Newman model of factors explaining health service utilization suggests three sets of independent variables [Anderson and Newman (1973)]. *Predisposing factors* include socio-demographic factors, here sex, race, and age. Each model also included an indicator of Medicaid blind/disabled eligibility. *Enabling factors* include measures of access to health care, both financial and geographic. All members of the sample had roughly equal financial access to care, given that they were dually eligible for Medicare and Medicaid. Direct measures of geographic access were not available, but each regression included the average value of the dependent variable during the baseline period for the individual's pharmacy cluster, effectively the individual's medical community defined via pharmacies with common prescribers. This variable served as a proxy for the net effect of geographic factors enhancing or reducing access to care and for geographic variation in the practice of care, such as variations in coding of conditions and diagnoses and in treatment protocols.

The third set of factors in the Andersen/Newman model are *measures of need*, such as diagnosis. For this study, the objective was to control for need, exclusive of the need that was influenced by the demonstration. The principal measure of need was based on the Diagnostic Cost Group (DCG) and Hierarchical Coexisting conditions (HCC) models and software developed by Ellis, Pope, Iezzoni, and others.[5] In short, the model used in this paper (the HCCPH-prospective model) reads information on a person's age, sex, diagnoses, life sustaining medical procedures, and selected hospitalizations during the baseline period and uses that information to predict Medicare payments for that person in the next year. As an additional measure of need, regressions controlled for the maximum number of distinct drugs, (as measured by NDC-9s), that the individual received during a single month of the baseline period.

The lengths of each person's follow-up and baseline periods were two other controls. Clearly, the longer the follow-up period, the greater that the recipient's expected health services utilization. The longer the baseline period, the more complete the record of the individual's prior health services utilization and, by proxy, ongoing medical problems. One would expect that the measure of expected health spending would be biased downward for some individuals with short baseline periods and therefore that, other things equal, a short baseline would be correlated with higher utilization in the follow-up period.

Finally, as participating pharmacies were recruited in waves rather than at a single point in time, each person was assigned an indicator that identified the recruitment wave of the pharmacy that filled that person's initial prescription. Hypothesizing a wave effect implies a belief in the possibility of a "halo" effect, in which a pharmacy's demonstration participation itself, independent of access to OPDUR, may affect behavior. Two competing hypotheses about how wave affects utilization might be entertained. On the one hand, late-wave treatment and control pharmacies might be better informed than early-wave pharmacies about the demonstration, leading to increased pharmacist CS activity and, presumably, reductions in downstream utilization in both arms. On the other hand, the time to observe utilization is foreshortened for late-wave recruits, suggesting that, even after adjustment for length of follow-up, utilization effects will be less frequent for late than for early wave recruits.

## Regression Specification

The model used to estimate demonstration effects was specified in the following way:

$$y_{ijt_f}=\alpha+\beta_1 tr_{ij}+X_i\beta_2+Z_i\beta_3+\beta_4 y_{jt_b}+\mu_j+\nu_{ij}$$

with:

$y_{ijtf}$ utilization of recipient i associated with cluster j during the follow-up period

---

[5] These models are described more fully in R. P. Ellis. G. C. Pope, L. Iezzoni et al, "Diagnosis-Based Risk Adjustment for Medicare Capitation Payments, *Health Care Financing Review*, Spring, 1996.

| | |
|---|---|
| $tr_{ij}$ | indicator that recipient i filled a prescription at treatment pharmacies and not at control pharmacies during the demonstration period |
| $X_i$ | predisposing (socio-demographic) and need characteristics of recipient i |
| $Z_i$ | length of follow-up and baseline windows for recipient i |
| $y_{jtb}$ | cluster average of outcome variable in baseline period |
| $\mu_j+\upsilon_{ij}$ | error |

The compound error term is the structure required for the variable effects model. As noted above, models using both the usual single-component error and compound error terms were tested, showing no differences in sign or significance of estimates of treatment effects. Findings below are based on the single-component error model.

Models predicting downstream utilization were estimated for:

- Component and total utilization measures, for each of the five populations/subpopulations;
- Selected probabilities (of inpatient admissions, of professional services, of all services) and total payments for all services, for services with diagnoses related to the misuse of each of seven drug categories (ACE inhibitors, antidepressants, antipsychotics, benzodiazepines, calcium channel blockers, digoxin, NSAIDs, total targeted)[6], for each of the five populations/subpopulations

### 5.4.3 Findings

*Descriptive Statistics*

Covariate means are useful: 1) for comparing the principal analysis population (dually eligible elders) with selected subpopulations and 2) for showing that, though pharmacy assignment in this demonstration was random (in clusters), randomization did not succeed in minimizing differences between treatment and control populations. Exhibit 5.4.4 shows the means of the variables included in the regressions, for all five analysis populations used in the study.[7] Subpopulations were slightly younger than the dually-eligible elderly population from which they were drawn; from 8 to 9 percent were 85 or over, compared to 13 percent in the elderly population. African Americans were more prevalent among elderly diabetics and the non-elderly. Gender differences between elderly (18 percent male) and non-elderly (49 percent male) are large and in the expected direction. Compared to the entire sample of elderly, men are less frequently represented in the diabetic, multiple prescription and prescriber samples. Expected eligibility differences between elderly and non-elderly reflect programmatic rules, not changes in acuity or chronicity of conditions with aging. Finally, predicted costs and use of multiple baseline drug measures clearly distinguish the greater health services needs

---

[6] Statistical power was insufficient to conduct tests of the effects of access to OPDUR on downstream utilization associated with H2 RA drugs.

[7] Certain qualitative variables excluded from the multivariate analyses are not shown: age (20-29) for the non-elderly, age (65-69) for the elderly, race (white) for all samples.

**Exhibit 5.4.4**

**HCFA DUR Demonstration Evaluation**
**Iowa OPDUR Demonstration Downstream Analysis Sample Characteristics**

| Measure | Population | | | | |
|---|---|---|---|---|---|
| | **All elderly** (n=4,597) | **Non-elderly** (n=4,478) | **Diabetics** (n=951) | **3+Prescribers** (n=990) | **10+Prescriptions** (n=489) |
| Age(30-39) | | -- | | 0.32 | ------ |
| Age(40-49) | | -- | | 0.22 | ------ |
| Age(50-59) | | -- | | 0.20 | ------ |
| Age(60-64) | | -- | | 0.02 | ------ |
| Age(70-79) | | 0.46 | | -- | 0.50<br>0.48<br>0.48 |
| Age(80-84) | | 0.16 | | -- | 0.13<br>0.16<br>0.16 |
| Age(85+) | | 0.13 | | -- | 0.08<br>0.08<br>0.09 |
| Race(black) | | 0.04 | | 0.07 | 0.07<br>0.05<br>0.05 |
| Race(other) | | 0.02 | | 0.01 | 0.02<br>0.01<br>0.01 |
| Sex(male) | | 0.18 | | 0.49 | 0.13<br>0.15<br>0.15 |
| Blind/disabled | | 0.11 | | 0.92 | 0.12<br>0.11<br>0.11 |
| DCG/HCC | | $4,247 | | $3,351 | $6,200<br>$5,593<br>$7,405 |
| Baseline drugs | | 2.1 | | 1.7 | 4.57.4<br>11.7 |

Source: Iowa OPDUR Demonstration Medicare/Medicaid Files\

of subpopulations. DCG/HCC predicts annual costs for the subpopulations from 32 to 74 percent above average costs for the whole elderly population. Subpopulation members use from two to five and a half times the number of drugs used in a month by the average elderly population member.

As has been stated, the design of the demonstration involved random assignment of geographic clusters of pharmacies to the intervention and control arms. Exhibit 5.4.5 shows the means of selected variables for intervention and control recipients separately. There were few significant differences between the two groups. Non-white persons (black and other racial minorities) were more prevalent in treatment than in control groups, and there was weak evidence that follow-up periods were slightly longer for treatment group members. The regression specifications controlled for these measured characteristics, and the variable effects model adjusts the standard errors to take account for any residual explanatory power associated with the cluster.

## Multivariate Analyses

### *Model Performance*

Though the analyses provided no consistent support for the hypothesis that access to OPDUR reduces downstream utilization, models used in the analyses generally performed well, given the limited availability of data for measuring covariates. Most models explained from 4 to 10 percent of variation in the dependent variables. Signs and significance levels of some model covariates were consistent across many of the models. For example, measures of baseline medical need, including predicted annual costs, maximum number of drugs in a month, and baseline measures of the dependent variables, were inevitably positive and highly significant. The number of follow-up months was also positive and significant, as expected. In addition, individuals who filled their first prescriptions in pharmacies recruited later into the study generally had lower probabilities and smaller levels of downstream utilization than , independent of the lengths of their follow-up periods. Demographic variables were less likely to generate significant estimates, but the estimated directions of effects (signs) were reasonably consistent for some measures. For example, African Americans were less likely than whites to experience an inpatient admission. Age measures generally depicted increasing frequency of utilization with advanced age, but decreasing intensity. Coefficients in a model predicting the costs of those individuals admitted for an inpatient stay were uniformly negative, and suggestive of some tendency for a reduction in costs for the oldest category.

### *Predictions of Total Utilization: All Study Populations*

CS payment has no consistent effects on components and aggregate measures of downstream utilization. Exhibit 5.4.6 shows the signs and levels of significance for treatment variables in models that predict inpatient, professional services, emergency room and total utilization in the five study populations/subpopulations. The exact numerical values of coefficients and their associated p-values provide no useful information, in the absence of any meaningful patterns in the estimates.

Few rows or columns in this exhibit show consistency of signs, and there are no instances in which the direction of effects, for any population or measure of utilization, is always negative. All but one coefficient for the non-elderly population are negative, but only one (the probability of an ER admission) is significant, at a 10 percent level. Though only six of forty coefficients are statistically

Exhibit 5.4.5

### HCFA DUR Demonstration Evaluation
### Iowa OPDUR Demonstration Downstream: Dually Eligible Elderly Treatment and Control Recipient Characteristics

| Measure | Treatment (N = 2,114) | Control (N = 2,483) |
|---|---|---|
| Age | 75.3 years | 75.6 years |
| Race(black)*** | 0.06 | 0.03 |
| Race(other)*** | 0.03 | 0.01 |
| Sex(male) | 0.18 | 0.18 |
| DCG/HCC | $5,325 | $5,484 |
| Baseline drugs | 4.9 | 4.8 |
| Baseline months | 10.8 | 10.9 |
| Follow-up months* | 17.0 | 16.5 |

* The difference between the arms is statistically significant at the 10 percent level.
**The difference between the arms is statistically significant at the 5 percent level.
***The difference between the arms is statistically significant at the 1 percent level.

Source: Iowa OPDUR Demonstration Medicare/Medicaid Files

Exhibit 5.4.6

## HCFA DUR Demonstration Evaluation Regression Adjusted Estimates of Total Utilization Effects: All Analysis Populations

| Utilization Measures | Populations | | | | |
|---|---|---|---|---|---|
| | All elderly | Non-elderly | Diabetics | 3+Prescribers | 10+ Prescription |
| Inpatient (0,1) | (-)* | (-) | (-)*** | (-) | (-) |
| Inpatient ($) | (+)** | (-) | (+) | (+)* | (+) |
| Professional services (0,1) | (-) | (+) | (-)*** | (-) | (-) |
| Professional services ($) | (-) | (-) | (+) | (-) | (-) |
| ER (0,1) | (+) | (-)* | (-) | (+) | (-) |
| ER ($) | (+) | (-) | (+) | (+) | (+) |
| Total (0,1) | (-) | (-) | (-)** | (-) | (-) |
| Total ($) | (-) | (-) | (-) | (+) | (-) |

* The difference between the arms is statistically significant at the 10 percent level.
** The difference between the arms is statistically significant at the 5 percent level.
***The difference between the arms is statistically significant at the 1 percent level.

Source: Iowa OPDUR Demonstration Medicare/Medicaid Files.

significant, in two populations, there are hints of underlying patterns. Both inpatient coefficients are negative and statistically significant for all elderly individuals. Professional and total services coefficients are also negative, but not significant, while ER coefficients are positive and insignificant. For the diabetic elderly, the probabilities of inpatient, professional services, and total services utilization are significantly lower for treatment group members. Given the ambiguity of the other estimates in this table, this finding suggests less an effect of the demonstration than a defect in the design, perhaps the influence of unmeasured characteristics of treatment individuals that make them more costly than controls to treat in an ER setting.

### *Predictions of Targeted Utilization: All Elderly, Elderly Diabetics, Elderly with 10 or More Prescriptions*

Despite what appears to be a repetition of the null findings reported above for total utilization, there is some evidence that, for some drug categories, treatment group members used fewer services than control group members . Exhibit 5.4.7 shows signs and significance levels for inpatient and professional services utilization associated with seven drug categories, separately and aggregated, for all dually eligible elderly, elderly diabetics and elderly with 10 or more prescriptions. Hypothesized negative signs on inpatient and professional utilization appear consistently across the three populations for calcium channel blockers and digoxin; these estimates are significant, however, only for elderly diabetics. Professional services utilization for conditions related to antidepressant and NSAIDs misuse appear to be reduced for treatment individuals, though the estimates are only significant for multiple drug users. Most strikingly, the probabilities of using any targeted professional services are significantly lower for treatment individuals in all three groups, and the signs on professional services payments are negative, though not significantly negative.

### *Comparing Demonstration Subjects with Screen Failures*

The Iowa demonstration recorded screen failures on prescriptions submitted by Medicaid customers of both treatment and control pharmacies. Since treatment pharmacies received messages back specifying the nature of the screen failure, and control pharmacies received no messages, it is reasonable to expect downstream utilization effects of OPDUR to be more prevalent in the subpopulation that generated screen failures. Based on simple comparisons of inpatient and professional services utilization (probability, total expenditure) between arms, restricting the comparisons to dually eligible individuals with one or more screen failures in the five analysis subpopulations, no patterns in signs of levels of statistical significance were detected. This was true, both for comparisons of all utilization and of utilization defined by diagnosis categories linked to the eight PSU drug groups.

### *Comparing Iowa and Washington Estimates of Effects: Targeted Utilization*

In general, comparisons of downstream effects between Iowa and Washington should be undertaken, and interpreted, with extreme caution. Despite many similarities in design and sampling methods, earlier chapters have shown the clear differences between these two states in baseline frequencies of prescription drug problems and drug utilization. The demonstrations chose interventions grounded in very different theories on what incentives motivate pharmacists' CS activities. Iowa OPDUR served many fewer dually eligible individuals than Project C.A.R.E., so that a comparable test for effects will, other things the same, be less precise in Iowa than in Washington. Nonetheless, since Iowa

**Exhibit 5.4.7**

**HCFA DUR Demonstration Evaluation**
**Regression Adjusted Estimates of Drug/Diagnosis Specific Utilization Effects: All Elderly, Elderly Diabetics, Elderly With 10+ Prescriptions**

| Drug Categories/ Utilization Measures | Populations | | |
|---|---|---|---|
| | All Elderly (N=2,898) | Elderly Diabetics (N=739) | Elderly/10+ Prescriptions (N=461) |
| *ACE Inhibitors* | | | |
| Inpatient (0,1) | (- ) | (-) | (+) |
| Professional services (0,1) | (+) | (-) | (-) |
| *Antidepressants* | | | |
| Inpatient (0,1) | (+) | (+) | (+) |
| Professional services (0,1) | (-) | (-) | (-)* |
| *Antipsychotics* | | | |
| Inpatient (0,1) | (-) | (+) | (-) |
| Professional services (0,1) | (-) | (+) | (-) |
| *Benzodiazepines* | | | |
| Inpatient (0,1) | (+) | -- | (+) |
| Professional services (0,1) | (-) | (+) | (-)* |
| *Calcium Channel Blockers* | | | |
| Inpatient (0,1) | (-) | (-)* | (-) |
| Professional services (0,1) | (-) | (-)* | (-) |
| *Digoxin* | | | |
| Inpatient (0,1) | (-) | (-) | (-) |
| Professional services (0,1) | (-) | (-) | (-) |
| *NSAIDs* | | | |
| Inpatient (0,1) | (+) | (-) | (+)* |
| Professional services (0,1) | (-) | (-) | (-)** |
| *All targeted (0,1)* | | | |
| Inpatient (0,1) | (-) | (-) | (-) |
| Professional services (0,1) | (-)** | (-)** | (-)*** |
| *All targeted ($)* | | | |
| Inpatient ($) | (+) | (+) | (+) |
| Professional services ($) | (-) | (-) | (-) |

*The difference between the arms is statistically significant at the 10 percent level.
**The difference between the arms is statistically significant at the 5 percent level.
***The difference between the arms is statistically significant at the 1 percent level.
Source: Iowa OPDUR Demonstration Medicare/Medicaid Files.

implemented an OPDUR screening system targeted at specific drug categories and criteria, whereas Washington did not target its CS payment demonstration, it would be reasonable to expect clearer evidence for the hypothesized effects on drug/diagnosis related utilization measures in Iowa than in Washington. To facilitate this comparison, the findings on targeted downstream effects reported in Exhibits 4.4.7 and 5.4.7 have been combined in Exhibit 5.4.8.

Though certain patterns emerge, it is difficult to find clear evidence that Iowa effects are stronger or more consistent than effects in Washington:

- In two of the three populations (all elderly and multiple drug users), negative signs (significant and insignificant) are more frequent in Iowa than in Washington (72 percent in both Iowa populations, compared to 50 percent and 61 percent in Washington elderly and multiple drug user populations respectively). The opposite is true for diabetics (83 percent in Washington, and 71 percent in Iowa).

- Iowa and Washington each produce eight statistically significant and negative estimates of treatment effects. Washington also produces three significant positive effects, against one for Iowa. Of the 17 "pairs" (by drug category/utilization measure and population) for which at least one estimate is significant, 10 agree on the appropriate sign. However, in only three instances (inpatient and professional services utilization associated with calcium channel blocker misuse, among elderly diabetics, and NSAIDs -related professional services use among multiple drug users) do both estimates agree (and, for NSAIDs, the estimates are both significant and opposite in sign).

- In both Iowa and Washington, association with treatment pharmacies seems to be related to reduced inpatient and professional services utilization linked to the misuse of calcium channel blockers. This pattern is strongest among elderly diabetics in both programs, though signs are uniformly negative throughout.

- Combining all targeted services, two patterns emerge. On the one hand, though estimated inpatient effects are negative in both states, estimates for diabetics and multiple drug users are statistically significant Washington but not for Iowa. On the other hand, in Iowa, the treatment group in all three populations uses significantly fewer targeted professional services than the control group, a pattern not replicated in Washington.

### 5.4.4 Discussion

The Iowa data provide no consistent evidence that OPDUR leads to significant changes in downstream utilization. Despite isolated "patterns" in utilization associated with the use of calcium channel blockers and digoxin, the overall picture is one of inconsistency and few instances of statistical significance. It had been expected that there would be evidence for effects within the utilization measures associated with the eight drug categories screened by Iowa OPDUR that was at least more compelling than evidence from Washington, which did not target specific drugs. This was not the case.

**Exhibit 5.4.8**

**HCFA DUR Demonstration Evaluation**
**Comparisons of Iowa and Washington Drug/Diagnosis Specific Utilization Effects: All Elderly, Elderly Diabetics, Elderly With 10+ Prescriptions**

| Drug Categories/ Utilization Measures | Populations | | | | | |
|---|---|---|---|---|---|---|
| | All Elderly | | Elderly Diabetics | | Elderly/10+ Prescriptions | |
| | IA | WA | IA | WA | IA | WA |
| *ACE Inhibitors* | | | | | | |
| Inpatient (0,1) | (-) | (+)** | (-) | (-) | (+) | (+) |
| Professional services (0,1) | (+) | (-) | (-) | (-) | (-) | (-) |
| *Antidepressants* | | | | | | |
| Inpatient (0,1) | (+) | (+) | (+) | (-) | (+) | (+) |
| Professional services (0,1) | (-) | (+) | (-) | (+) | (-)* | (+) |
| *Antipsychotics* | | | | | | |
| Inpatient (0,1) | (-) | (-) | (+) | (-) | (-) | (-) |
| Professional services (0,1) | (-) | (-) | (+) | (-)* | (-) | (-) |
| *Benzodiazepines* | | | | | | |
| Inpatient (0,1) | (+) | (+) | -- | (+) | (+) | (-) |
| Professional services (0,1) | (-) | (+) | (+) | (-) | (-)* | (+) |
| *Calcium Channel Blockers* | | | | | | |
| Inpatient (0,1) | (-) | (-) | (-)* | (-)** | (-) | (-)** |
| Professional services (0,1) | (-) | (-) | (-)* | (-)* | (-) | (-) |
| *Digoxin* | | | | | | |
| Inpatient (0,1) | (-) | (-)* | (-) | (-) | (-) | (-) |
| Professional services (0,1) | (-) | (+)** | (-) | (-) | (-) | (-)* |
| *NSAIDs* | | | | | | |
| Inpatient (0,1) | (+) | (-) | (-) | (-) | (+)* | (+) |
| Professional services (0,1) | (-) | (+) | (-) | (+) | (-)** | (+)** |
| *All targeted (0,1)* | | | | | | |
| Inpatient (0,1) | (-) | (-) | (-) | (-)** | (-) | (-)* |
| Professional services (0,1) | (-)** | (+) | (-)** | (-) | (-)*** | (-) |
| *All targeted ($)* | | | | | | |
| Inpatient ($) | (+) | (- ) | (+) | (-) | (+) | (+) |
| Professional services ($) | (-) | (+) | (-) | (-) | (-) | (-) |

* The difference between the arms is statistically significant at the 10 percent level.
** The difference between the arms is statistically significant at the 5 percent level.
***The difference between the arms is statistically significant at the 1 percent level.
Source: Iowa OPDUR Demonstration Medicare/Medicaid Files.

## 5.5 Effects of the OPDUR Demonstration on Prescription Drug Use and Costs

This section presents findings of the effect of the Iowa OPDUR demonstration on prescription drug use and costs. The objective of this part of the evaluation was to determine whether there were statistically significant reductions in drug cost and use between Medicaid recipients who filled their prescriptions in demonstration pharmacies relative to pharmacies in the control group during the demonstration period. The hypothesis is that by alerting pharmacists to potential prescribing problems before they occur, OPDUR screening will lead to improvement in drug therapy, lower drug usage rates, and lower cost in those therapeutic classes subject to review. Since the Iowa demonstration was targeted to the eight therapeutic classes reviewed by the PSU screener, the primary focus of the analysis is upon those drug groups. It is also conceivable that DUR screening will have a halo effect and lead to improved prescribing in other therapeutic domains. To test that hypothesis, the analysis focused on all drugs combined and on the same nine high-expense drug groups evaluated with respect to the Washington C.A.R.E. project (see Section 4.5).

It is recognized that—irrespective of OPDUR—most Iowa pharmacies utilized some type of proprietary DUR screening software during the demonstration period. The presence of such software reduces the expected demonstration effect size in all of the recipient groups studied. However, because in-store DUR software packages typically review patient drug histories for prescriptions filled only at that store, they can miss problems that arise when patients use multiple prescribers and multiple pharmacies. The Iowa OPDUR system, with its central repository of all Medicaid prescriptions should enable pharmacists to expand their reach to catch such problems. The study subjects most likely to be affected are those with three or more prescribers. Consequently, we hypothesize stronger OPDUR treatment effects for this group compared to the others.

The study design closely follows that presented in the previous section on downstream medical utilization and costs. For this reason, the discussion of data and methods is quite brief consisting primarily of a summary of study procedures. No strong evidence of a demonstration effect was found in any of the five subpopulations reviewed. Regression results are presented for the dual-eligible 65+ age group and for elderly with three or more prescribers. Results from the other study groups are briefly discussed.

### 5.5.1 Data and Study Population

This aspect of the study utilized the same analytic dataset as for the downstream analyses reported in Section 5.4. The study population was also the same, consisting of dual-eligible (Medicare/Medicaid) Medicaid recipients who met the following inclusion and exclusion criteria: (1) eligible for both Medicare and Medicaid during the baseline and follow-up periods, (2) exposed to the demonstration, and (3) did not fill prescriptions in both a demonstration and control pharmacy (see Exhibit 5.4.1). Application of these criteria resulted in a sample of 4,597 elderly subjects and 4,478non-elderly subjects. The elderly sample was further subset into three at-risk groups: elderly diabetics (951), elderly with three or more prescribers (990) and elderly with 10 or more prescriptions (489). The sample selection process is described more fully in Section 5.4.2.

### 5.5.2 Methods

The statistical methods employed to test the study hypotheses are identical to those used in the downstream analyses for both Washington and Iowa (see Sections 4.4.2 and 5.4.2). Multiple regression models were specified to produce conditional estimates of the demonstration impact on a wide range of drug use and cost outcomes holding constant other factors that might contaminate or bias the demonstration coefficients. The dependent variables in these models included total prescriptions, total Medicaid drug payments, individual totals for each of the eight PSU drug classes (ACE inhibitors, calcium channel blockers, digoxin, antidepressants, antipsychotics, benzodiazepines, H2RAs, and NSAIDs), and individual totals for nine high-expense drug categories (central nervous system agents, gastrointestinal drugs, cardiovascular drugs, analgesics and anesthetics, endocrine and metabolic drugs, respiratory agents, anti-infectives, neuromuscular drugs, and topical products). In all, 270 multiple regression models were estimated.

Except for the choice of dependent variables, model specifications were identical to those used in the downstream analyses. The independent variables are listed in Exhibit 5.4.3. Demonstration effects on the probability of any prescription drug use in the follow-up period were estimated with Probit. Single-component error OLS models were used to estimate demonstration effects on counts and costs of drugs used by those subjects with at least one prescription fill in the follow-up period. Because the distributions of drug cost and use are right-skewed, the dependent variables in these regressions were transformed into log form. The coefficients in these models can thus be read as elasticities.

Test regressions with interaction terms involving the demonstration effect variable failed to produce significant results. The same was true for regressions estimated using a variable effects specification. These results are not reported.

### 5.5.3 Findings

The demographic characteristics of the five study populations are summarized in Exhibits 5.4.4 and 5.4.5 in the downstream analysis section. As can be seen there, the study subjects are predominantly white females in both entitlement categories (aged and blind/disabled), but there are other major differences between the two groups. Individuals entitled through disability or blindness are much younger (modal age, 30-39), much more likely to be male (49% versus 18% for the aged), and have lower predicted resource use according to the DCG/HCC index ($3,351 versus $4,247). The "at risk" elderly subpopulations are somewhat younger but much sicker than the elderly sample as a whole (the DCG/HCC index for elderly with 10+ prescriptions is 75 percent above that for the whole). As noted in the previous section, the control and treatment groups differ significantly in selected domains (age, race, and baseline drug use) thereby necessitating the statistical control offered by multiple regression techniques.

Exhibit 5.5.1 presents means and standard deviations for the dependent variables on behalf of the 65+ dual-eligible sample. The top panel of the table reports utilization and cost statistics for the nine high-expense drug groups. The bottom panel reports utilization and cost in the eight PSU screener drug categories. Totals for all drug use are shown in the bottom row. By definition, all subjects filled at least one prescription in the follow-up period (evidence of drug use was necessary for assigning subjects to the treatment or control group). The average elderly recipient utilized 30 prescriptions or refills during the 13 months of the follow-up period at an average cost of $498.

All of the elderly at-risk subpopulations had much higher usage rates. Elderly diabetics filled an average of 52 scripts in the follow-up period. Recipients with three or more prescribers filled 67 prescriptions, while those with 10 or more prescriptions filled an average of 84. Average drug costs were also much higher for these subpopulations, ranging from $916 for elderly diabetics to $1,510 for the 10+ prescription group. Dual-eligible recipients under age 65 averaged 25 prescriptions at a cost of $464 during the follow-up period.

The distribution of drug use by therapeutic class shown in Exhibit 5.5.1 followed an expected pattern. The elderly used more cardiovascular drugs than any other type of prescription medication (77 percent of the sample used these drugs at a mean cost of $228 in the follow-up period). More than 60 percent of the sample filled one or more prescriptions for analgesics/anesthetics and anti-infectives, and except for neuromuscular drugs, more than 39 percent filled prescriptions in the other high-expense drug classes. Among the PSU screener drugs, utilization rates were highest for calcium channel blockers, NSAIDs, and H2RAs. Relatively few elderly Iowa recipients took antipsychotics (7 percent) or benzodiazepines (5 percent).

**Exhibit 5.5.1**

**HCFA DUR Demonstration Evaluation**
**Characteristics of Iowa Study Sample Recipients' Drug Utilization by Therapeutic Class: Dual-eligible 65+ (n=4,597)**

| | Sample Statistics | | | | | |
|---|---|---|---|---|---|---|
| | Percent Using Drugs | | Count of Drugs Used | | Cost of Drugs Used | |
| Drug category (high expense) | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| Central nervous system agents | 43.3 | 0.50 | 5.9 | 3.5 | $52.45 | $5.93 |
| Gastrointestinal agents | 41.2 | 0.49 | 5.0 | 6.6 | 142.59 | 6.17 |
| Cardiovascular drugs | 77.2 | 0.42 | 13.6 | 3.6 | 228.15 | 5.81 |
| Analgesics and anesthetics | 62.2 | 0.49 | 5.1 | 3.4 | 38.09 | 5.87 |
| Endocrine and metabolic drugs | 39.7 | 0.49 | 6.9 | 3.4 | 92.76 | 5.58 |
| Respiratory agents | 37.6 | 0.48 | 4.4 | 3.8 | 69.41 | 6.11 |
| Anti-infectives | 64.6 | 0.48 | 2.9 | 2.5 | 40.04 | 4.31 |
| Neuromuscular drugs | 12.4 | 0.33 | 3.7 | 3.3 | 55.70 | 6.11 |
| Topical products | 43.2 | 0.50 | 2.9 | 2.7 | 43.82 | 4.22 |
| Drug category (PSU classes) | | | | | | |
| ACE Inhibitors | 23.0% | 0.42 | 6.2 | 2.9 | $208.51 | $3.86 |
| Antidepressants | 21.3 | 0.41 | 4.8 | 3.2 | 56.83 | 5.70 |
| Antipsychotics | 7.3 | 0.26 | 2.7 | 3.1 | 28.79 | 4.76 |
| Benzodiazepines | 5.4 | 0.23 | 3.6 | 3.2 | 22.65 | 3.94 |
| Calcium channel blockers | 32.1 | 0.47 | 7.5 | 3.0 | 262.43 | 4.35 |
| Digoxins | 20.1 | 0.40 | 5.8 | 2.9 | 35.87 | 3.32 |
| H2RAs | 26.5 | 0.44 | 4.6 | 3.1 | 249.64 | 3.82 |
| NSAIDs | 26.5 | 0.44 | 3.6 | 2.9 | 45.15 | 5.75 |
| All drugs | 1.00 | 0.00 | 30.0 | 4.3 | $497.70 | 6.89 |

Source: Medicaid data from Iowa, 1994-96.

Exhibit 5.5.2 presents multivariate results of the hypothesis test that the Iowa OPDUR project reduced drug use and cost among the elderly dual-eligible sample. The cells in this table report coefficient values and standard errors for the treatment/control dummy variable from the 30 regressions relating to total drug use and drug use within the PSU screener drug. It is clear from this table that there was no treatment effect. Only one coefficient (Medicaid payments for digoxin) reached conventional levels of statistical significance. Despite the fact that 25 of the 30 coefficients were in the hypothesized direction (-), a statistical test for randomness of signs (the Wilcoxon-Mann-Whitney rank sum test) could not reject the null hypothesis of random (+) and (-) signs. Regression results from the nine high-expense drug groups produced similar findings—mostly negative coefficients but none significant (the Wilcoxon-Mann-Whitney test could not rule out randomness in signs here either).

**Exhibit 5.5.2**

**HCFA DUR Demonstration Evaluation**
**Regression Results of the Impact of the Iowa OPDUR Demonstration on the Probability of any Drug Use, Counts of Drugs, and Cost of Drugs Used in the Eight PSU Drug Classes, Dual Eligibles Aged 65+**

| | Regression Coefficients (standard Errors) for the Demonstration Effect | | | | | |
|---|---|---|---|---|---|---|
| | Probability of Use (Probit) | | Count of Drugs (ln) | | Cost of Drugs (ln) | |
| PSU Drug Category | | | | | | |
| Ace inhibitors | 0.00 | (0.04) | -0.08 | (0.05) | -0.10 | (0.07) |
| Antidepressants | -0.02 | (0.04) | -0.11 | (0.07) | -0.12 | (0.11) |
| Antipsychotics | 0.01 | (0.06) | -0.10 | (0.12) | -0.25 | (0.16) |
| Benzodiazepines | -0.08 | (0.07) | -0.02 | (0.13) | -0.14 | (0.16) |
| Calcium Channel Blockers | -0.05 | (0.04) | -0.04 | (0.04) | -0.05 | (0.06) |
| Digoxin | -0.03 | (0.05) | -0.06 | (0.05) | **-0.16 | (0.06) |
| H2 Receptor Antagonists | -0.03 | (0.04) | -0.04 | (0.06) | -0.04 | (0.07) |
| NSAIDs | 0.02 | (0.04) | 0.04 | (0.06) | 0.12 | (0.10) |
| Drugs in all eight classes | -0.02 | (0.04) | -0.03 | (0.03) | -0.04 | (0.05) |

Source: Medicaid drug data from Washington, 1993-1996.
Notes: Statistical significance indicated by * if p<.05 and ** if p<.01 in two-tailed tests.

The regression equations for the at-risk elderly subpopulations and for the under 65 dual-eligible sample also failed to generate any evidence of treatment effects. Regression results for the subpopulation of elderly with three or more prescribers are shown in Exhibit 5.5.3. The pattern of nonsignificant negative coefficient values is evident in these results. As hypothesized, the coefficient values are mostly negative and they are uniformly higher than in the regressions for the elderly population as a whole. However, only two coefficients reached conventional levels of statistical significance. Application of the Wilcoxon-Mann-Whitney test could not rule out randomness in signs.

As in the case of the downstream analyses, the regression models themselves performed relatively well in terms of overall explained variance. The coefficients of determination (R2) in the OLS regressions (adjusted for number of covariates) ranged from 6 to 51 percent in the therapeutic class regressions and between 58 and 85 percent in the total drug use regressions. As expected, the two variables with the greatest predictive power were the number of follow-up months (the longer the observation period the greater the use and cost) and the maximum number of GPI classes represented in the baseline period (also positive). Other explanatory variables that generally reached statistical significance were the

HCC/DCG cost predictor (+), sex (women use more drugs), other race (persons other than white or black use fewer drugs), and age (-). These findings are similar to those reported in the downstream analyses (Section 5.4.3) and the explanations offered there apply here as well.

**Exhibit 5.5.3**

**HCFA DUR Demonstration Evaluation**
**Regression Results of the Impact of the Iowa OPDUR Demonstration on the Probability of any Drug Use, Counts of Drugs, and Cost of Drugs Used in the Eight PSU Drug Classes, Dual Eligibles Aged 65 + with 3 or More Prescribers**

| | Regression Coefficients (standard Errors) for the Demonstration Effect | | | | | |
|---|---|---|---|---|---|---|
| | Probability of Use (Probit) | | Count of Drugs (ln) | | Cost of Drugs (ln) | |
| PSU Drug Category | 0.05 | (0.09) | -0.02 | (0.10) | -0.05 | (0.13) |
| Ace inhibitors | -0.03 | (0.09) | *-0.23 | (0.11) | -0.27 | (0.18) |
| Antidepressants | 0.10 | (0.11) | -0.39 | (0.21) | -0.40 | (0.29) |
| Antipsychotics | -0.13 | (0.12) | 0.08 | (0.23) | -0.11 | (0.26) |
| Benzodiazepines | -0.05 | (0.09) | -0.13 | (0.08) | -0.15 | (0.11) |
| Calcium Channel Blockers | -0.12 | (0.10) | -0.01 | (0.08) | -0.17 | (0.10) |
| Digoxin | -0.05 | (0.09) | -0.15 | (0.10) | -0.16 | (0.13) |
| H2 Receptor Antagonists | 0.06 | (0.09) | -0.01 | (0.11) | 0.13 | (0.18) |
| NSAIDs | -0.03 | (0.13) | *-0.14 | (0.06) | -0.14 | (0.09) |
| Drugs in all eight classes | | | | | | |

Source: Medicaid drug data from Washington, 1993-1996.
Notes: Statistical significance indicated by * if p<.05 and ** if p<.01 in two-tailed tests.

## 5.5.4 Discussion

The analysis presented in this section offers no support for the hypothesis that the Iowa OPDUR demonstration had a measurable impact on either drug use or cost. The regression results fail to show any pattern of downward trends attributable to the demonstration in any of the five population groups for any of the 18 drug classes studied.

It is possible that different results would have been obtained had the study focused on different population groups (e.g., women and children enrolled in AFDC) or other drug classes. But this seems very unlikely. Dual-eligible recipients tend to be among the sickest members of the Medicaid population and the highest users of prescription drugs. Were OPDUR alerts to have any impact on drug use and cost, one would expect to find it in this group of recipients. Failure to find an impact implies either that the OPDUR alerts are ineffectual in influencing pharmacist behavior or that the marginal impact of the State's prospective DUR system was lost in the background noise of competing in-store DUR systems. Unfortunately, there is no obvious analytic way to distinguish between these two conclusions. The IOWA OPDUR demonstration was designed expressly to evaluate the impact prospective DUR, yet it was implemented in a real-world setting in which in-store DUR systems are ubiquitous. Partialling out the independent effects of each is virtually impossible. The lack of a contemporaneous comparison series also makes it impossible to formally assess the combined impact of all the DUR systems in place at any given time, whether in Iowa or elsewhere.

## 5.6 Effects of the OPDUR Demonstration Project on Screen Failures

The Iowa OPDUR Demonstration Project was funded by the Health Care Financing Administration as a demonstration of on-line prospective drug utilization review. By using computerized algorithms to screen prescriptions at the point-of-sale for therapeutic problems, this prospective DUR system has significant potential to reduce drug therapy problems.

The Iowa OPDUR system utilizes information about prescriptions filled at any pharmacy in the state that accepts Medicaid; this feature of the system allows the pharmacist to make more informed decisions than would be possible with a standard pharmacy computer system. Furthermore, the Iowa OPDUR algorithms are based on the University of Maryland and Philadelphia College of Pharmacy and Sciences (UM/PCPS) criteria, which constitute the most refined set of criteria currently available. It was hypothesized therefore that the OPDUR system in Iowa would be effective in preventing drug therapy problems which have the potential to result in costly and potentially serious adverse health outcomes.

Other chapters of this report have examined the impact of the Iowa OPDUR Demonstration Project on drug utilization, medical expenditures, and adverse clinical outcomes. The goal of this chapter is to assess the impact of the Iowa OPDUR Demonstration Project on the incidence of screen failures. To focus on screen failure rates is reasonable because if clinical improvements (e.g. reduction in rates of adverse outcomes) occurred, they would be expected to be paralleled by or preceded by reductions in screen failure rates. Specifically, if the OPDUR system was effective in improving prescribing, then rates of screen failures would be expected to decline in the treatment group throughout the follow-up period.

As this chapter will show, with a few exceptions, the treatment and control group patients generally did not have statistically significant differences in screen failure rates. The analysis presented in this chapter therefore suggests that the Iowa OPDUR system was limited in its ability to effect prescribing changes.

### Evaluation Approach

The objective of the present study was to examine the association between the Iowa OPDUR system and the occurrence of screen failures associated with eight drug classes. The expectation was that pharmacists using the OPDUR system would be more likely than pharmacists not using the OPDUR system to detect prescription problems and to consult with prescribers regarding therapeutic alternatives. Over time, as prescribers had increased contact with pharmacists and had opportunities to modify their prescribing, the incidence of new screen failures was therefore expected to decline.

To address these issues, a multivariate analytic design was employed. For the screen failure outcome, defined as any screen failure during the follow-up period, the residual effect of the demonstration was examined after adjustment for known risk factors. Multivariate regression procedures were used to control for demographic variables, drug utilization, and other characteristics that could affect a given patient's likelihood of screen failures. All control variables were measured during a baseline period prior to the onset of the demonstration so as to be independent of any demonstration effects.

### Data

The data used in this study were extracted from Medicaid drug claims and Medicare files for patients in the state of Iowa in 1993 (to provide a baseline) and 1994-1996. OPDUR was in operation in Iowa between June 1994 and June 1997. Other records and data streams generated by the demonstration were also incorporated into the data set. Data development is described in detail in Chapter 5.7.

*Data Base and File Structure*

Pharmacies enrolled in the Iowa OPDUR demonstration in two waves (June 1994 and December 1994). Although other approaches are possible, this study assumes that a patient's exposure to the demonstration (and consequently the patient's follow-up period) began on the day that the person first went to an active demonstration pharmacy and lasted until December 1996, loss of eligibility for either Medicaid or Medicare, entry into a nursing home, or death, whichever came first. Similarly, baseline periods ended the day before the person's first contact with a demonstration pharmacy and began twelve months prior to the ending date or on the date that the patient first became eligible for the Medicare program, became eligible for the Medicaid program, left an HMO, or left a nursing home and became a resident in the community. These windows were constructed in order to capture distinct time periods during which the demonstration could and could not have influenced screen failure rates.

*Study Populations*

The main population for this study included Iowa residents who were for Medicaid and who filled prescriptions at pharmacies participating in the OPDUR demonstration. For the analyses presented in this chapter, additional requirements were that participants must: 1) have three or more months of baseline data; 2) have an adequate follow-up period; 3) not be enrolled in a managed care program; and 4) not reside in a nursing home.

The final population of eligible study participants included individuals both younger than and older than age 65. For the present study, the elderly represent a population of particular interest because on average the elderly experience poorer health and receive more prescription drugs than younger individuals. These factors, along with aging-related alterations in pharmacokinetics, place the elderly at much greater risk for numerous adverse drug reactions. It was expected that the elderly should represent a subgroup to which the OPDUR intervention would most likely be targeted and therefore a subgroup that would most likely experience a decline in screen failures. Moreover, even among the elderly there are subgroups whose drug therapy is highly likely to present problems: elderly receiving prescription from multiple prescribers, and elderly receiving a large number of prescription within a relatively short time period. Lastly, it was thought that diabetics, because of the nature of their disease condition and/or drug therapy were expected to benefit from OPDUR.

Exhibit 5.6.1 outlines the selection of the final study population and subsamples from the original pool of Medicaid eligible persons who filled prescriptions at any demonstration pharmacy between 1993 and 1997 (N=145,764). A total of 66,412 individuals were eliminated due to incomplete baseline or follow-up data. An additional 13,286 individuals were eliminated because they either had incomplete demographic data or because they utilized pharmacies in both arms (control and demonstration) of the study.

**Exhibit 5.6.1**

**HCFA DUR Demonstration Evaluation**
**Derivation of Iowa OPDUR Screen Failure Analysis Samples**

| | |
|---|---|
| Medicaid eligible, filled prescriptions at demonstration pharmacy | 145,764 |
| *Less:* | |
| Individuals with insufficient data (no follow-up period, baseline less than three months) | (66,412) |
| Not exposed due to timing of first prescription | (959) |
| No demographic data available | (4,376) |
| Filled prescription in both arms of study | (7,951) |
| Net usable study population | 66,066 |
| Dually eligible (Medicaid/Medicare) | 9,707 |
| Non-elderly (< 65 at end of study) | 4,478 |
| Became 65 during study (omitted from sample) | (632) |
| Elderly (65+ at beginning of study, 9/30/95) | 4,597 |
| *Subpopulations of elderly eligibles:* | |
| Elderly diabetic | 951 |
| Elderly with 3+ prescribers in any baseline month | 990 |
| Elderly with 10+ prescriptions in any baseline month | 489 |

Source: Iowa OPDUR Demonstration Project Medicare/Medicaid Files

The final study sample of 9,075 dually-eligible recipients included similar numbers of individuals under 65 (N=4,478) and those 65 or older (N=4,597). The elderly population (N=4,597) was further subset into groups especially likely to benefit from the OPDUR demonstration. These included:

1) Elderly diabetics – individuals who had claims for insulin or oral hypoglycemic agents at any point during 1993-1995 (N=951).

2) Elderly with three or more prescribers during any baseline month (N=990)

3) Eligible elderly with 10 or more prescriptions during any single baseline month (N=489).

### Dependent Variables

All of the outcomes selected for this portion of the study involved the occurrence of screen failures during the follow-up period. Downstream screen failure measures included indicators both of whether or not a screen failure occurred, and, for those with any screen failures, how many failures were generated during the follow-up period.

The DUR criteria screened for this study were originally developed by expert panels convened by the University of Maryland and the Philadelphia College of Pharmacy and Science (UM/PCPS) under a cooperative agreement with HCFA (Knapp and Erwin, 1992). The screening process involved a review of all Medicaid prescriptions within the following drug categories: ACE inhibitors, antidepressants, antipsychotics, benzodiazepines, calcium channel blockers, digoxin, $H_2$-receptor antagonists, and non-steroidal anti-inflammatory agents. The types of drug problems screened included four prescribing problems -- dosage, drug/drug interactions, duration of therapy, and therapeutic duplication.

Screening of claims to identify prescriptions which violated the established DUR criteria was conducted using the DUR screener developed at Penn State (the PSU screener). For the present study, which examined the effect of a prospective online DUR system, it was critical for the DUR screener to be able to replicate the screening being conducted online by OPDUR. For purposes of the present study, then, the PSU screener was modified to duplicate the specific edits used in the OPDUR implementation. Comparisons of baseline and follow-up screen failure rates could then be accomplished by using the modified PSU screener to screen all claims both prior to and following the introduction of OPDUR. Details regarding the PSU screener and the 61 DUR criteria used in this study have been provided in Chapters 2 and 3 of this report.

All analyses employed both broadly defined measures, such as any'screen failure for any drug, and also more narrowly-defined measures involving either specific therapeutic classes or specific DUR criteria. In order to examine therapeutic class differences in screen failure rates, analyses were conducted for the eight drug classes represented in the PSU screener criteria and for the aggregate variable representing any screen failure over the total 61 screens contained in all eight therapeutic classes. In addition to the analyses at the level of therapeutic class, other analyses examined screen failure rates within 30 of the 61 specific screens that are generated by the PSU screener. These 30 screens were selected based on an observed annual prevalence greater than 5% in any of the four states for which screen prevalence rates were computed[8].

### Independent Variables

In modeling the effects of the demonstration on subsequent screen failures, measures were first developed to characterize demonstration exposure, then, based on the available data, additional covariates were identified which were hypothesized to impact on screen failure rates independently from the effects of the demonstration. The independent variables selected for these analyses are listed in Exhibit 5.6.2 and are briefly detailed below.

---

[1] Three screens were omitted from this rule because of low denominators: Nimodipine duration, Ketrolac with probenecid, and Mefenamic acid or phenybutazone duration.

*Exposure to the Intervention*

Although the OPDUR demonstration was designed as a random experiment, randomization occurred at the level of the pharmacy cluster while the impact of the intervention is measures at the level of the patient. This posed a significant methodological problem, because individuals had the potential to fill prescriptions at different pharmacies and, potentially, to fill prescriptions at a mix of intervention, control, and non-demonstration pharmacies. It could be expected that the effect of the demonstration would be diluted for such individuals depending on the relative proportions of prescriptions filled at the different pharmacy types. To resolve this problem, this study adopted a simple approach. Individuals were considered members of the intervention group if they ever had a Medicaid-paid drug claim at an intervention pharmacy and never had such a claim at a control pharmacy. Similarly, they were considered members of the control group if they ever had a Medicaid-paid drug claim at a control pharmacy and never had such a claim at an intervention pharmacy. Individuals who went to both intervention and control pharmacies were excluded from the study regardless of the number of filled prescription at either intervention or control pharmacies[9].

*Other Covariates*

***General Predisposing and Need Characteristics***. The basic analytic strategy for the several clinical outcomes studies reported here was to examine the effect of the treatment variable within a regression framework that statistically controlled for demographic factors and other measures thought to be related to the risk of invoking a screen failure. For all outcomes, the regression models included variables for the patient's race, sex, and age. Additionally, a dichotomous variable was created to indicate whether or not a person's Medicaid eligibility was associated with being blind or disabled. Another variable was created to control for medical need, independent of the need that was influenced by this demonstration. The principal measure of need was based on the Diagnostic Cost Group (DCG) and Hierarchical Coexisting conditions (HCC) models and software developed by Ellis et al. (1996). The DCG/HCC model uses information on a person's age, sex, diagnoses, life-sustaining medical procedures, and selected hospitalizations during the baseline period and uses that information to predict Medicare payments for that person in the next year. As an additional measure of need, regressions controlled for the maximum number of distinct drugs that each individual received during a single month of the baseline period. These factors could affect screen failure rates indirectly as proxies for health status. They may also affect the probability of having access to, or using, health care services. Thus these measures serve to control for general predisposing and need characteristics.

***Rate of Screen Failure in Pharmacy Cluster.*** The pharmacy cluster, the unit of randomization in this analysis, effectively defined medical communities. Potentially, there could be unmeasured differences in demographic and behavioral characteristics among patients in different pharmacy clusters that influenced the need for care and the predisposition to seek it. Similarly, there could be geographic variation associated with cluster in access to care and in its practice, such as variations in coding of conditions and diagnoses and in treatment protocols. If these differences were associated with measured clinical outcomes, then controlling for them would improve the precision of estimated intervention effects. Thus, each model included the average screen failure rate during the baseline period for each pharmacy cluster as an additional regressor. This variable served as a proxy for the net effect of geographic factors enhancing or reducing access to care and for geographic variation in the practice of care, such as variations in prescribing patterns for specific conditions.

---

[2] For this purpose, the day that the patient first went to the demonstration pharmacy was included in the follow-up period.

*Length of Baseline and Follow-up Periods*. It would be expected that patients with longer follow-up periods would be more likely to experience one or more screen failures than those with shorter follow-up periods. Therefore, the length of the follow-up period was included as a covariate in all models. While the length of the baseline period does not directly affect patients' probabilities of generating screen failures, it is likely that the assessment of drug utilization and other risk factors may be measured with less precision for patients with short baseline periods relative to patients with full baseline periods. Thus, controlling for other measured covariates, the probability of screen failures might vary according to the length of the baseline period; consequently, the length of the baseline period was included as an additional covariate. For all models, the lengths of the follow-up and baseline period were entered as the natural log of the number of months in each period.

*Pharmacy Wave*. Pharmacies entered the cognitive services demonstration in two waves. The first wave of 119 pharmacies enrolled in June 1994. A second wave of 104 pharmacies of pharmacies entered in December 1994. It is possible that pharmacies entering the demonstration later learned from the experiences of prior entrants and used the OPDUR system more effectively. For these reasons, indicator variables were included in the models to capture the wave of the first demonstration pharmacy where the individual filled a prescription. For all models, the reference category was wave one. The basic set of independent variables that were used in every model (as described above) are summarized below in Exhibit 5.6.2.

**Exhibit 5.6.2**

**HCFA DUR Demonstration Evaluation**
**Independent Variables: Iowa OPDUR Downstream Screen Failure Analysis**

| Variable | Definition |
|---|---|
| *Demonstration exposure* | |
| Treatment (0,1) | Demonstration status of individual: treatment if all prescriptions during study period filled at treatment pharmacies, control if all filled at control pharmacies |
| *Other independent variables* | |
| Age | Nine qualitative variables: 20-29, 30-39, 40-49, 50-59, 60-64, 65-69, 70-79, 80-84, 85+ |
| Sex | Two qualitative variables: male, female |
| Race | Three qualitative variables: white, black, other |
| Eligibility | Medicaid eligibility - blind or disabled |
| DCG/HCC | Predicted annual Medicare payments (proxy measure of medical need) |
| Baseline screen failures | Pharmacy cluster average of screen failures during the baseline period |
| Baseline drugs | Maximum number of distinct drugs in a single month during the baseline period; natural logarithm of value entered into regression models |

**Exhibit 5.6.2**

**HCFA DUR Demonstration Evaluation**
**Independent Variables: Iowa OPDUR Downstream Screen Failure Analysis**

| Variable | Definition |
|---|---|
| Length of baseline period | Measured in months; natural logarithm of value entered into regression models |
| Length of follow-up period | Measured in months; natural logarithm of value entered into regression models |
| Wave | Wave of pharmacy entrance (for pharmacy filling individual's first prescription) into demonstration (two waves in Iowa) |

## Models and Statistical Methods

Because the dependent variables denoting any screen failure were categorical and binary, probit regressions were used. The regression specification for the examination of treatment effects in this study was:

$$y^{c}_{ijt_f}=\alpha+\beta_1 tr_{ij}+X_{it_b}\beta_2+W^{c}_{it_b}\beta_3+Z_i\beta_4+\beta_5 y^{c}_{jt_b}+\nu_{ij}$$

with:

| | |
|---|---|
| $y^{c}_{ijtf}$ | indicator that patient i associated with pharmacy cluster j experienced screen failure c during the follow-up period (tf) |
| $tr_{ij}$ | indicator that patient i filled a prescription at treatment pharmacies and not at control pharmacies during the demonstration period |
| $X_{it}$ | general predisposing (socio-demographic) and need characteristics of patient i determined during the baseline period |
| $W^{c}_{itb}$ | patient i's risk factors for screen failure c determined during the baseline period |
| $Z_i$ | length of follow-up and baseline windows for patient i |
| $y^{c}_{jtb}$ | rate of screen failure c in pharmacy cluster j during the baseline period |
| $\nu_{ij}$ | error |

## Findings

### *Descriptive Statistics*

Exhibit 5.6.3 below displays the means of the independent variables included in the regressions, for all five analysis populations used in the study. Some differences between the elderly and non-elderly groups primarily reflect program rules. For example, the non-elderly group is mainly comprised of blind or disabled participants (92%), compared to 11% in the elderly group. Elderly recipients in general had higher expected medical expenses and took more baseline drugs than non-elderly. Relative to the elderly population, a much higher percentage of the non-elderly were male (49% versus 18%). The three elderly subpopulations differ from the total elderly population in several respects: not surprisingly, diabetic or high-utilization elderly had higher expected medical costs and used more drugs than the overall elderly group. The five populations do not differ greatly in the proportion of individuals visiting pharmacies in the treatment groups or pharmacies enrolling in the demonstration during Wave 2.

**Exhibit 5.6.3**

**HCFA DUR Demonstration Evaluation**
**Iowa OPDUR Screen Failure Analysis Sample Characteristics**

| | Population | | | | |
|---|---|---|---|---|---|
| Measure | Non-elderly (N=4,478) | All elderly (N=4,597) | Diabetic elderly (N=951) | Elderly with 3+ prescribers (N=990) | Elderly with 10+ Rx's |
| Age (30-39) | 0.32 | | | | |
| Age (40-49) | 0.22 | | | | |
| Age (50-59) | 0.20 | | | | |
| Age (60-69) | 0.02 | | | | |
| Age (70-79) | | 0.46 | 0.50 | 0.48 | 0.48 |
| Age (80-84) | | 0.16 | 0.13 | 0.16 | 0.16 |
| Age (85+) | | 0.13 | 0.08 | 0.08 | 0.09 |
| Race (black) | 0.07 | 0.04 | 0.07 | 0.05 | 0.05 |
| Race (other) | 0.01 | 0.02 | 0.02 | 0.01 | 0.01 |
| Sex (male) | 0.49 | 0.18 | 0.13 | 0.15 | 0.15 |
| Blind or disabled | 0.92 | 0.11 | 0.12 | 0.11 | 0.11 |
| Wave 2 | 0.45 | 0.44 | 0.46 | 0.40 | 0.45 |
| Treatment group | 0.50 | 0.46 | 0.47 | 0.47 | 0.45 |
| DCG/HCC | $3,351 | $4,243 | $6,200 | $5,593 | $7,405 |
| Baseline drugs | 1.7 | 2.1 | 4.5 | 2.0 | 2.5 |
| Baseline length (months) | 10.2 | 10.3 | 10.7 | 11.5 | 11.0 |
| Follow-up length (months) | 12.3 | 11.0 | 12.4 | 14.2 | 11.0 |

Source: Iowa OPDUR Project Medicare/Medicaid Files

*Treatment Effects*

It was hypothesized that, relative to individuals visiting control pharmacies, individuals who visited pharmacies participating in the OPDUR system would be less likely to experience screen failures during the follow-up period. If the feedback to pharmacists provided by the OPDUR system did in fact significantly decrease subsequent prescribing problems likely to generate OPDUR alerts, then the models should indicate that, after adjusting for baseline demographic and risk factors and for factors related to potential design effects, the treatment group should have had lower rates of screen failures. If this were the case, the probit coefficients should be negative and significant. Similarly, among groups of individuals who experienced any screen failure during the follow-up period, the OLS coefficients obtained from predicting total number of screen failures from the treatment variable should also be negative and significant.

***Effects within eight therapeutic classes***. Exhibit 5.6.4 provides the coefficient signs and their respective statistical levels of significance for the demonstration treatment effects within the eight therapeutic drug classes defined by the PSU screener. Results for all probit models are presented; results for OLS models are presented for those outcomes for which OLS models were run. (OLS models were only performed for subpopulations or drug categories in which the total number of persons experiencing screen failure was 100 or greater.)

**Exhibit 5.6.4**

## HCFA Demonstration Evaluation
## Regression Adjusted Estimates of Treatment Effects: All Analysis Populations

| Therapeutic Category | Non-elderly | All Elderly | Elderly diabetics | Elderly with 3+ prescribers | Elderly with 10+ prescriptions |
|---|---|---|---|---|---|
| ACE Inhibitors | | | | | |
| Any failure | (-) | (+) | (+) | (+) | (+) |
| Number of failures | (-)* | (-) | ---- | (+)* | ---- |
| Antidepressants | | | | | |
| Any failure | (+) | (-) | (-) | (-) | (-) |
| Number of failures | (+) | (+)* | ---- | ---- | ---- |
| Antipsychotics | | | | | |
| Any failure | (+) | (+) | (+) | (+) | (-) |
| Number of failures | (+)* | (-) | ---- | ---- | ---- |
| Benzodiazepines | | | | | |
| Any failure | (-) | (-) | (+) | (-)* | (+) |
| Number of failures | (+) | (-) | ---- | (+) | ---- |
| Calcium channel blockers | | | | | |
| Any failure | (+) | (-) | (-) | (-) | (+) |
| Number of failures | (-) | (-) | ---- | ---- | ---- |
| Digoxin | | | | | |
| Any failure | (-) | (+) | (-) | (+) | (+) |
| Number of failures | (-) | (-) | ---- | ---- | ---- |
| $H_2$RA's | | | | | |
| Any failure | (-) | (+) | (+) | (-) | (_) |
| Number of failures | (+) | (-) | ---- | ---- | ---- |
| NSAID's | | | | | |
| Any failure | (-) | (-) | (-) | (-) | (-)* |
| Number of failures | (-) | (-) | ---- | (-) | ---- |
| All Drug Classes | | | | | |
| Any failure | (+)* | (-) | (-)* | (-) | (-) |
| Number of failures | (+) | (-) | (+) | (+) | (+) |

| | |
|---|---|
| ---- | OLS not performed (N < 100) |
| * | The difference between the arms is statistically significant at the 0.10 level |
| ** | The difference between the arms is statistically significant at the 0.05 level |
| *** | The difference between the arms is statistically significant at the 0.01 level |

Examination of the model results summarized in Exhibit 5.6.4 suggests that the Iowa OPDUR system was not greatly effective in preventing and reducing screen failures during the follow-up period. For the probit models, which evaluated the effect of treatment on any screen failure while controlling for other variables, 26 out of the 45 estimated models (58%) yielded treatment coefficients that were negative in sign. Only three of these coefficients, however, were statistically significant at a level of 0.10. These included: 1) for elderly diabetics, the outcome of any failure across all eight classes (p=0.080); 2) for elderly with 3 or more prescribers, the outcome of any benzodiazepine failure (p=0.085); and 3) for elderly with 10 or more prescriptions, the outcome of any NSAID failure (p=0.063). In addition to these findings, one probit coefficient for the treatment effect was positive and significant; this effect occurred in the non-elderly subpopulation for the outcome of any screen failure across all eight drug classes (p=0.081).

In addition to the probit models, OLS regression models were also performed which examined the association between the independent variables and the number of screen failures generated during the follow-up period. In order to ensure the minimal sample size necessary for model estimation, these OLS regressions were only performed within sub-populations and drug classes in which the total number of individuals experiencing one or more screen failures exceeded 100. The available OLS results, which are also presented in Exhibit 5.6.4, do not suggest any clear effects of the treatment exposure on level of screen failures among individuals experiencing any screen failure. Approximately half of the OLS regression coefficients for the treatment effect variables were negative in sign. Four of the treatment coefficients were significant at 0.10, including three positive coefficients: 1) non-elderly, antipsychotic failures (p=0.065); 2) total elderly, antidepressant failures (p=0.096); and 3) elderly with three or more prescribers, any ACE inhibitor (p=0.084). Only one statistically-significant coefficient was negative in sign: non-elderly, ACE inhibitor failures (p=0.063).

The positive associations of treatment with screen failures is unexpected. However, it is unlikely that these positive effects of treatment on screen failure level are indicative of any systematic response to the OPDUR system. In order for systematic positive associations to have occurred, either pharmacists or prescribers would have to have responded unfavorably to the OPDUR intervention and in a manner that increased subsequent screen failure rates. It seems unlikely, even if prescribers disagreed with or responded negatively to the OPDUR system, that they would actively increase their rate of prescribing of regimens that would fail the OPDUR screener.

***Effects within specific screening criteria***. In addition to examining the treatment effects by broad therapeutic class, examination of the treatment effects for specific criteria may also yield useful information regarding the effectiveness of the OPDUR system in reducing screen failures. Exhibit 5.6.5 provides the probit coefficient signs and their respective statistical levels of significance for the demonstration treatment effects for 30 specific screen criteria having failure rates high enough to warrant further analyses[10].

---

3 Despite requiring the cutoff for the prevalence rate for screen failures to be higher than 5%, there were a number of instances where the probit or OLS analyses could not be conducted due to low occurrences of the screen failures.

**Exhibit 5.6.5**

## HCFA Demonstration Evaluation
## Regression Adjusted Estimates of Treatment Effects: 30 Specific Screening Criteria

| Screen Class and Name | Non-elderly | All elderly | Elderly diabetic | Elderly 3+ pre-scribers | Elderly 10+ Rx's |
|---|---|---|---|---|---|
| ACE Inhibitors | | | | | |
| Adult dosage | (+) | (+) | (+) | (-) | (-) |
| Multiple ACE inhibitors | (-) | (-) | (+) | (-) | (-) |
| Antidepressants | | | | | |
| Multiple antidepressants | (-) | (-) | (-) | (-) | (-) |
| Tricyclic antidepressant or maprotiline with clonidine | (-) | (+) | (-) | (-) | na |
| Antipsychotics | | | | | |
| Adult dosage | (+)*** | (-)** | (-) | (-) | (-) |
| Antipsychotic with antiparkinsonian | (+) | (-) | na | (-) | na |
| Multiple antipsychotics | (+) | (+) | (+) | (+) | (-) |
| Benzodiazepines | | | | | |
| Adult dosage | (-) | (-) | (+) | (+) | (-) |
| Benzodiazepines, duration | (+) | (-) | (+) | (-) | (-) |
| Benzodiazepine with antihistamine | (-) | (-)** | (+) | (-) | (-) |
| Benzodiazepine with anxiolytic | (-) | (-) | (+) | (+) | (-) |
| Benzodiazepine with sedative/hypnotic | (-) | (-) | (+) | (-) | (-) |
| Estazepam, flurazepam, temazepam, or triazolam, duration | (-) | (-)* | (+) | (-) | (-) |
| Long half-life benzodiazepine over age 65 | na | (-) | (-) | (-) | (-) |
| Multiple benzodiazepines | (-)* | (-) | (+) | (-) | (+) |
| Calcium Channel Blockers | | | | | |
| Adult dosage | (+) | (-) | (-) | (-) | (-) |
| Multiple calcium channel blockers | (-) | (-) | (-) | (-) | (+) |
| Verapamil with beta blocker | (+) | (+) | (-) | (+) | (+) |
| Digoxin | | | | | |
| Adult dosage | (-) | (-) | (+) | (+) | (+) |

**Exhibit 5.6.5**

**HCFA Demonstration Evaluation**
**Regression Adjusted Estimates of Treatment Effects: 30 Specific Screening Criteria**

| Screen Class and Name | Non-elderly | All elderly | Elderly diabetic | Elderly 3+ prescribers | Elderly 10+ Rx's |
|---|---|---|---|---|---|
| $H_2$RA's | | | | | |
| Adult dosage | (+) | (-) | (-) | (-) | (-) |
| Cimetidine with phenytoin, methylxanthines, or warfarin | (+) | (+) | (+) | (-) | (+) |
| $H_2$RA acute duration | (+) | (+) | (+) | (+) | (+) |
| $H_2$RA and sucralfate on same day | (-) | (-) | (-) | (-) | (+) |
| $H_2$RA GERD duration | (-) | (+) | (+) | (-) | (-) |
| $H_2$RA with misoprostol, omeprazole, or sucralfate | (-)** | (+) | (-) | (-) | (-) |
| Multiple $H_2$RA's | (-) | (-) | (+) | (-) | (-) |
| NSAID's | | | | | |
| Adult dosage | (-) | (-) | (-) | (-) | (-)* |
| Mefenamic acid or phenylbutazone duration | na | na | na | na | na |
| Multiple NSAID's | (-) | (+) | (+) | (-) | (+) |
| Salicylates with oral antidiabetic | (-) | (-) | (-) | (+) | (+) |

| | |
|---|---|
| na | The number of screen failures was either zero or too small for model convergence |
| * | The difference between the arms is statistically significant at the 0.10 level |
| ** | The difference between the arms is statistically significant at the 0.05 level |
| *** | The difference between the arms is statistically significant at the 0.01 level |

Despite requiring the cutoff for the prevalence rate for screen failures to be higher than 5%, there were a number of instances where the probit or OLS analyses could not be conducted due to low occurrences of the screen failures.

Overall, the results shown in Exhibits 5.6.4 and 5.6.5 indicate little treatment effect on the occurrence of screen failures within the eight drug classes screened and for the total across all drug classes. No discernable patterns involving specific patient sub-populations or therapeutic classes are apparent. These findings suggest that the OPDUR system was relatively ineffective in changing the prescribing patterns which led to the OPDUR alerts within and across these therapeutic classes.

### *Other Covariates*

Race and gender generally were not strong predictors of screen failure in the populations studied here, although there were a few exceptions. For example, in the non-elderly, black race was associated with lower rates of screen failures for benzodiazepines ($p < 0.01$) and NSAID's ($p < 0.01$). For the non-elderly, male sex was strongly associated with lower rates of screen failures within the broad therapeutic classes of antidepressants ($P < 0.001$) and NSAID's ($p < 0.001$). Not surprisingly, length of follow-up time was consistently associated with a higher probability of screen failure. In general, the wave of pharmacy entrance into the demonstration was not associated with the likelihood of failing the screen criteria. In most study groups and therapeutic classes, measures of medical need (such as predicted Medicare expenditures) and drug utilization (such as maximum number of drugs used during a baseline month) were strongly associated with higher rates of screen failure.

### *Pharmacy Cluster Effects*

To account for the fact that clusters of pharmacies rather than individuals were randomized to the treatment and control groups, each model included as a covariate the average rate of baseline screen failure in the individual's pharmacy cluster. It was expected that this variable would help control for unobservable, cluster level differences in factors that affected the frequency of screen failures. Across the five populations and eight drug classes screened, these variables were consistently and very strongly associated with screen failure.

## Comment

The Iowa OPDUR system was designed to alert pharmacists to prescription problems at the point of service so that inappropriate prescribing could be reduced. This chapter focused on the impact of the intervention on the incidence of screen failures. The analysis controlled for treatment and comparison group characteristics that may have differed at baseline, as well as the cluster-based sampling design used in the demonstration. Based on the regression results discussed above, there is little evidence that the OPDUR messages prevented subsequent screen failures.

There are a number of reasons that may explain why few significant effects were found. Because previously-validated criteria were used in the OPDUR system, it is unlikely that the OPDUR system failed to identify potentially serious drug therapy problems. However, in order for a reduction in risk to occur, a specific chain of events had to take place once a screen failure was generated. The success of the OPDUR system in reducing inappropriate therapy is largely dependent upon several key components. These include the actions of the pharmacists receiving the alerts, the actions of the prescribing physicians, and the behavior of patients. The first and most immediate factor is the pharmacist who receives the screen failure alert. Upon receiving an alert, the pharmacist has the option of overriding the alert and dispensing the medication as written, or contacting the prescribing physician to inform them of the therapeutic issue flagged by the OPDUR system. The attitudes and behaviors of the prescribing physician are a second critical component influencing the outcome of OPDUR alerts. In general, reduction in therapeutic problems and adverse outcomes associated with screen failures will be limited primarily by the willingness and ability of physicians to modify their prescribing patterns. Without additional data indicating what specific actions were taken by pharmacists and prescribers in response to the OPDUR messages, it is difficult to draw any strong conclusions regarding which system components were particularly ineffective.

It is also likely that limitations in the study design complicated the measurement and detection of existing treatment effects, particularly if the treatment effects were very small. Generally, a cluster sampling approach reduces the power of detecting differences between the treatment and control groups. In this case, however, the intracluster correlation is likely to be low, and many of the clusters were very small; cluster based sampling therefore does not seriously reduce the power of the study design. More important is the fact that, if the treatment effect is expected to be very small then the study had limited power to detect a differences. To capture small differences between treatment and control groups, larger sample sizes would be needed. Although these sample size constraints greatly limited the ability of the study to detect significant effects, the findings nevertheless suggest that the Iowa OPDUR demonstration did not lead to large declines in screen failure rates.

## 5.7 Outcomes Analyses – Methods and Data

This report presents findings on the impact of the Iowa OPDUR system on adverse outcomes related to use of four classes of prescription drugs — NSAIDS, antidepressants, cardiovascular drugs, and benzodiazepines. These four studies are presented in separate chapters, and each study used methods that were designed to appropriately address the specific clinical issues that were of interest. Nevertheless, all of the outcomes studies were based on the same data and variable creation methodology. Furthermore, to a large extent, the statistical models used in the four studies were very similar. The objective of this chapter is to present an overview of the data, general evaluation approach and variable creation methods that were used across all of the outcomes studies. Each outcomes study chapter summarizes this general approach and also describes methods and variables used that were specific to the study.

### General Evaluation Approach

The objective of each outcomes study was to explore the relationship between OPDUR and the rate of adverse events related to use of a particular set of prescription drugs. It was hypothesized that pharmacists using the OPDUR system would be more likely than pharmacists not using the OPDUR system to detect prescription problems related to prescription drug use and, consequently, to have an opportunity to prevent adverse outcomes. Patients who visited pharmacies where prescriptions were screened by OPDUR, therefore, should be less likely than control group patients to experience illnesses related to inappropriate prescribing.

The Iowa OPDUR demonstration randomly assigned pharmacy clusters to treament and control groups. The treatment pharmacies received OPDUR messages while the control group pharmacies had prescriptions screened for research purposes but did not receive OPDUR messages. In the analysis, individuals were considered to have received treatment if they filled a Medicaid paid prescription at a treatment pharmacy at least once during the study period. Individuals were considered to be members of the control group if they filled a Medicaid paid prescription at a control pharmacy at least once during the study period. To avoid problems with crossover, individuals who filled Medicaid paid prescriptions at both treatment and control pharmacies were dropped from the analysis.

Each of the outcomes studies essentially compared the treatment and control groups' rates of clinical outcomes. Because clusters were randomized instead of individuals, there were some demographic differences between the treatment and control groups. For this reason, multiple regression was used to assess the effect of the treatment on individuals' outcomes. The models controlled for demographic, disease, and drug factors predisposing a given patient to the outcome under study. All control variables were measured during a baseline period before the demonstration began so as to be independent of any demonstration effects. The treatment was measured using an indicator variable. Because all dependent variables were binary, probit regressions were used.

Thus, the regression specification was:

$$y_{ijt_f}^{c}=\alpha+\beta_1 tr_{ij}+X_{it_b}\beta_2+W_{it_b}^{c}\beta_3+Z_i\beta_4+\beta_5 y_{jt_b}^{c}+\nu_{ij}$$

with:

| | |
|---|---|
| $y^c_{ijtf}$ | indicator that patient i associated with pharmacy cluster j experienced clinical outcome c during the follow-up period (tf) |
| $tr_{ij}$ | indicator that patient i filled a prescription at treatment pharmacies and not at control pharmacies during the demonstration period |
| $X_{it}$ | general predisposing (socio-demographic) and need characteristics of patient i determined during the baseline period |
| $W^c_{itb}$ | patient i's risk factors for outcome c determined during the baseline period |
| $Z_i$ | length of follow-up and baseline windows for patient i |
| $y^c_{jtb}$ | rate of clinical outcome c in pharmacy cluster j during the baseline period |
| $\nu_{ij}$ | error |

### *Covariates*

Each outcomes study used as covariates a distinct set of variables that were expected to influence the outcomes under investigation. These variables are described in the outcomes study chapters that follow this chapter. A basic set of covariates, however, was used in every model across the four studies. In addition to the treatment dummy variable, this set included race, sex, age, length of baseline period, length of follow-up period, and the wave when the pharmacy entered the demonstration.

### *Correction for Cluster Sampling*

The fact that pharmacy clusters rather than individual patients were randomized required that the standard regression framework be adjusted. If indeed there were factors correlated with cluster that affected outcomes and if, despite the controls included in the models, these factors were not fully measured, then, because the number of clusters was small (191) relative to the number of patients, these differences might not average out between intervention and control groups, leading to misleading estimates of intervention effects. Specifically, traditionally-estimated standard errors would be smaller than the true standard errors.

The first line of defense against this problem was the inclusion of a full set of covariates, including a variable that measured the cluster average of the dependent variable. This cluster average was expected to capture factors that vary by cluster that are important in the determination of clinical outcomes. In addition, a cluster adjusted version of White's method was used to estimate the variance-covariance matrices of the models. This method relaxes the assumption of independence within clusters, and is considered to be a robust variance estimator in cases of cluster-based sampling.[11] An alternative approach would have been to bootstrap the coefficients of each model and derive nonparametric standard errors for the coefficients.[12] In two of the outcomes studies (antidepressant and cardiovascular drugs), this alternative method was tested; each anti-depressant and cardiac model was bootstrapped and standard errors were computed. The results, which are not presented in this report, were very similar to the results derived using the cluster-adjusted version of White's method.[13]

### *Maryland and Georgia Data*

During the study design phase of the evaluation, it was acknowledged that the outcomes analyses might be seriously limited by a lack of adequate power to detect differences in outcomes between the treatment and control groups. The treatment effect was expected to be quite small, some of the outcomes were very rare, and the number of dually eligible beneficiaries in Iowa was limited. For these reasons, steps were taken to increase the sample size in Iowa in the hopes that the power to detect differences could be enhanced.

Two states — Maryland and Georgia — were identified as possible proxies for the treatment and control groups. Because Maryland had an OPDUR program in operation during the study period, individuals from Maryland could be added to the treatment group. Individuals from Georgia could be added to the control group since Georgia never had an OPDUR system in operation. The Iowa analyses, therefore, could be replicated with these expanded treatment and control groups.[14]

---

[11] William Rogers originally developed this method in the early 1980's to account for cluster-based sampling issues in the RAND Health Insurance Experiment.

[12] The bootstrapping method involves pulling repeated samples (with replacement), estimating coefficients repeatedly (100 times in this case), and then deriving standard errors from the distributions of estimated coefficients. This method is often used in cases where we have no formulas to compute standard errors or when the assumptions we need to make to compute standard errors do not make sense (i.e. in this case, we could not assume that error terms were independently distributed because of inter-cluster correlation).

[13] Two other methods were considered to account for the effects of clustered sampling. The random effects probit was considered, but the large size of the clusters made this model very difficult if not impossible to implement computationally. Generalized Estimating Equations (GEE) methods could have been used, but this approach would have made estimation more complex and the advantages of this approach were uncertain.

[14] The number of dually eligible beneficiaries who filled prescriptions during the study period was very large in Maryland and Georgia. For this reason, the samples pulled from these two states were limited to users of the four deug classes that were of interest.

Undoubtedly, this approach has limitations. This method essentially merges two different study designs (a randomized controlled trial with volunteering pharmacies and a quasi-experimental design with all pharmacies) and it is not clear how the multiple design affects the interpretation of results. The addition of Maryland and Georgia introduces the problem of state-level counfounding factors affecting the outcomes of interest. Even so, it was hoped that the results of an analysis that includes Maryland and Georgia might confirm suggestive results from an analysis of Iowa only, which might be severely limited by low power.

The analysis of Iowa alone, however, did not lead to suggestive results. Instead, as the following four chapters describe in detail, the findings offer little if any evidence that the OPDUR system was effective in preventing outcomes related to inappropriate prescribing of NSAIDS, antidepressants, cardiovascular drugs, and benzodiazepines. Consequently, an analysis that included Maryland and Georgia is not presented in the main text of the chapters.

## Data and Variable Creation

The analysis of selected health outcomes among users of demonstration pharmacies is based on reviews of data derived from Medicaid and Medicare administrative sources. Administrative data is collected for the purposes of program management and to enable and record the payment of medical providers for services and supplies. The data is not collected for health research purposes, however there are a number of data fields that contain patient clinical and demographic information that can be used for the analysis of health outcomes. The data used in the analyses are demographic and program participation information from enrollment data and drug use, diagnosis and procedure information from medical claims billing records. The data available for the analyses was complete state Medicaid and Medicare data for services delivered in the period 1993-1996.

### *Enrollment Data*

The populations potentially affected by demonstration influences are Medicaid recipients eligible for pharmacy benefits. A significant portion of the Medicaid population are also entitled to benefits under the Medicare program. Enrollment data from both programs contains person-level information that details by date enrollment status. Included in these records are data describing personal demographic characteristics, managed care affiliation and the specific categories under which eligibility has been granted. From an epidemiological perspective the enrollment data is the source of all population denominator information. Population studies that focus on disease prevalence and the frequency of acute events require program eligibility data to define and model study groups. The models require the demographic information but also need to account for individuals who are not eligible over the study period and whose claims history is incomplete.

Most of the outcome analyses focused on the community dwelling elderly. This population is generally covered under Medicaid for pharmacy benefits and Medicare for institutional and physician services. Medicaid recipients who are not entitled to Medicare benefits have complete coverage under the Medicaid program for all medical products and services. To fully model the population the Medicaid and Medicare enrollment data were linked using common identifiers. The cross-program linkages were validated by ensuring matches based on coded identifiers as well as personal characteristics such as birthdate and gender. The integration of the enrollment data allowed for the monthly determination of eligibility status in both programs. The preparation of a joined Medicaid and Medicare enrollment file entailed the assignment of unique study identifiers to all records pertaining to an individual. Enrollment data is used in the analyses to identify eligible populations, determine baseline and follow-up observation windows and to assign values to the demographic covariates.

## *Pharmacy Claims*

Medicaid pharmacy claims are used in the outcome analyses to both identify the populations directly affected by the demonstration and to document drugs used by the patients in the baseline and follow-up periods. A pharmacy claim record reliably includes the following data elements: patient identifier, pharmacy identifier, date of dispense, drug entity and quantity dispensed. Prescription duration information is reported in pharmacy claim records but it is of variable quality. The outcome analyses uses drug claims data to identify Medicaid recipients who used demonstration pharmacies. Drug claims are also used to identify population risk groups based on pre-demonstration exposures to targeted therapeutic entities. Drug claims data served as a source for proxy information on diagnostic status when the presence of a therapy was a strong indicator of disease.

### *Physician Service Claims*

Medicaid and Medicare physician claims are used in the outcome analyses as source of diagnostic information and to detect the occurrence of targeted medical procedures. Physician claim records reliably report the following data elements: patient identifier, provider identifier, dates of service (at least to the month level), medical procedures performed and diagnoses. Diagnostic and procedure information are used in the analyses to determine outcome status in the follow-up period and for the formulation of baseline clinical covariate values. Two of the outcome analyses used the medical procedure codes in combination with the diagnosis codes to obtain a more secure outcome definitions. The analysis of hip fractures relied on a combination of diagnosis and a hip fracture repair procedure to identify an outcome and the analysis of gastrointestinal bleeding used as an outcome definitions an endoscopy procedure in combination with a GI bleed diagnosis. Other outcome analyses used only diagnosis codes and the measured effects are explicitly defined in terms of rates of reported symptomologies.

### *Institutional Service Claims*

Medicaid and Medicare institutional claims are used in the outcome analyses as sources of diagnostic and procedure information. Institutional claims can be for inpatient stays or outpatient visits. The claims are generally not directly for medical procedures but for use of institutional facilities, for example inpatient accommodations or use of treatment rooms. The following key data elements are present on institutional claims: patient identifier, provider identifier, dates of service, diagnoses, procedure and/or revenue center code. Institutional claims for inpatient and outpatient services are used in the analyses in the same way as physician claims: for the formulation of covariate values and the determination of outcome status. Outpatient institutional claims are especially useful as a data source for care delivered by medical residents who are not allowed to bill directly for their services; in this context no physician claim is submitted but the institution does bill for facility use and supports the claim with diagnosis and procedure information. Inpatient claims distinguish primary diagnoses from other diagnoses reported in the course of the hospital stay. This distinction can be important when an outcome identifies an acute event as opposed to a simple report of disease presence. One outcome analysis, gastrointestinal bleeding, used the primary diagnosis as a flag for a severe outcome event.

### *Analytic File and Variable Construction*

The enrollment and claims databases are organized into a relational database keyed by individual patient identifiers. The outcome analyses were based on person summary records constructed from the main databases. The content of the person summary records varies for each outcome analysis in the covariate and outcome information captured, however the underlying design is the same across analyses. Baseline and follow-up periods were calculated using eligibility data from the enrollment file and demonstration contact data derived from drug claims. Outcome indicators were established through reviews of diagnostic and procedure data reported on claims with service dates in the follow-up period. Clinical covariate indicators were established through reviews of drug use, diagnostic and procedure data reported on claims with service dates in the baseline period. The person summary files were designed for easy analysis using standard computerized statistical processing tools.

## 5.8 The Impact of OPDUR on Gastrointestinal Events

The HCFA DUR Demonstrations are designed to enhance Medicaid pharmacy practice through encouraging pharmacists to increase their review of patient drug therapies. The Iowa OPDUR Demonstration Project has focused on the pharmacist as a key provider in the final review of prescription orders. In order to enhance the tools available to pharmacists the OPDUR project has provider point-of-sale information systems to pharmacists on an experimental basis. The systems were linked in real-time to the Iowa Medicaid drug claims payment databases and were designed to review prescriptions prior to dispensing in the context of a patient's drug utilization history. The systems screen prescriptions for dosage violations, therapeutic duplication and drug-drug interactions using information gathered from all pharmacies used by the patient. Pharmacists participating in OPDUR were given training on system operations and the types screening rules that were embedded in the program. Pharmacies in the treatment arm were given systems that provided warning messages when screening violations occurred, pharmacists in the control group were provided with systems that only provided eligibility verifications, but did not transmit clinical warnings. All participating pharmacists also received training in the provision and documentation of patient counseling.

It was hypothesized that through providing pharmacists with computerized utilization review technology, the effectiveness of pharmacists as medical providers will be enhanced through wider and easier access to patient prescription information. Greater patient knowledge will result in a more through identification of patients who are vulnerable to prescription-related adverse events. The vulnerable populations served by Medicaid, which include indigent elderly people and children, are at higher risk for these prescribing problems and their associated adverse outcomes. These groups therefore are the most likely to benefit from an effective computerized utilization review program. The overall impact of the programs was anticipated to be the more efficient and therapeutic utilization of prescription drugs and, potentially, a decrease in health outcomes linked to avoidable adverse effects from pharmaceutical exposures. The division of participating pharmacies into treatment and control groups leads directly to an experimental design to test the impact of Medicaid computerized drug utilization review systems on health outcomes.

### Therapy Related GI Ulcer Formation and Hemorrhage

Gastrointestinal symptoms are widely associated with exposures to pharmaceutical agents. The most severe outcomes are hospitalizations for upper GI ulcer formation and bleeding linked to the use of Non-Steroidal Anti-Inflammatory Drugs (NSAIDs). Non-aspirin NSAIDs are one of the most frequently prescribed drug categories and aspirin is a common over the counter medication. The high frequency of NSAID use, the severity of the health effects and the wide dissemination in the medical community of information about the association of NSAIDs to GI events, all suggest that pharmacy level interventions may be effective in reducing GI outcomes.

It is widely recognized that the risk of gastrointestinal ulceration and hemorrhaging (GI bleed) is significantly enhanced by exposure to aspirin and non-aspirin Non-Steroidal Anti-Inflammatory Drugs. Since the 1980s the medical literature has included numerous papers documenting the relationship between exposures of NSAIDs to hospitalization for GI bleeding, a selective search of Medline for articles on GI outcomes related to NSAID exposure yielded 44 major new studies and comprehensive reviews in the last 10 years in the major medical journals. A review of population cohort studies which evaluated the risk of GI events as stratified by NSAID exposure showed odds ratios ranging from 1.4-2.1 [Willet et. al. (1994)]. The risk of GI outcomes is elevated in populations which have histories of GI disease, other co-morbidities that relate to GI symptoms or pharmacy therapies that increase the risk of GI bleeding. Several treatment modalities have been developed to reduce the frequency of NSAID related GI disease in at risk populations, including the substitution of non-NSAID agents, reduction in NSAID dosage levels and prophylactic co-therapies. Pharmacist based interventions that result in greater supervision of therapies may lead to reductions in the frequency of GI outcomes in a treated population.

Clinical and epidemiological studies of the relationship between NSAIDs and GI bleeding have used a variety of measures to determine the extent of the damage caused by the counter-prostaglandin action of NSAIDs. The studies have reviewed data ranging from survey data on GI symptoms to endoscopic evidence of tissue damage to mortality associated with severe episodes of GI bleeding. The population studies in the literature that use insurance claims data focus on hospitalization events and death, although several studies relied on outpatient diagnoses as well [Carson et. al. (1987)]. Due to the relative rarity of hospitalization for ulcer and hemorrhage there are four outcome types that are being studied with varying levels of specificity, severity and statistical power:

The analysis of the demonstration's impact on the incidence of GI bleeding was based on the premise that improved pharmacy practices can be manifested as both short and long term effects through the modification of one or more prescription regimens and review of over the counter drug use. The underlying mechanism for reduction in outcome risk could be increased coordination between patients, pharmacists and prescribing physicians leading to more effective supervision of prescription drug regimens. This analysis focuses on the question of whether the use of Iowa State OPDUR pharmacies that are reimbursed for computerized utilization review was related to decreases in GI events and symptomology in client populations. The analysis measures the risk of a GI outcome in populations that are exposed to pharmacies that receive messages from computerized utilization review systems in comparison to populations exposed to pharmacies who receive no messages.

## Analytic Methods

The analyses measured the relationship between use of a demonstration pharmacy and the risk of a patient outcome. The models employed in these measurements rely on drug utilization and diagnostic information derived from Medicaid and Medicare pharmacy and medical claims. The conclusions of the analyses are adjusted for the demographic and disease covariates listed as relevant in the medical literature. In total the outcome analyses present three specific assessments of the efficacy of the OPDUR program in improving the health of their patients according as measured by the odds ratio for GI outcomes in the pharmacy study groups.

- Physician or hospital diagnosis of an upper GI ulcer or hemorrhage;
- Physician or hospital diagnosis of an upper GI ulcer or hemorrhage and a GI endoscopy;
- Inpatient hospital diagnosis of upper GI ulcer or hemorrhage;
  - Inpatient primary hospital diagnosis of upper GI ulcer or hemorrhage;
  - Inpatient primary hospital diagnosis of upper GI ulcer or hemorrhage with subsequent death.

The objective of the analysis was to explore the relationship between the presence of computerized utilization review and the rate of GI related adverse events. It was hypothesized that computerized utilization review would increase the surveillance of prescription dosage and duration and therapeutic duplication. This increase, in turn, was expected to change prescriber and/or patient behavior, and make prescriptions safer and more effective. Multiple logistic[15] regression was used in order to control for demographic, disease, and drug factors predisposing a given patient to the outcome under study. All control variables were measured during a baseline period before the demonstration began so as to be independent of any demonstration effects.

### *Study Population*

The analytic unit of observation was the Medicaid recipient who had contact with a demonstration pharmacy. Person level analyses are called for since the risk of outcome is primarily determined by individual disease and exposure covariates. Because the coverage of Medicaid managed care has been increasing over the period of the demonstration there will be portions of the population who will not be analyzable because of a lack of drug exposure and diagnostic data. Additionally Medicaid recipients who are institutionalized will not be directly subject to OPDUR influences and will not be included in the models. As a result the analyses will focus on the non-nursing home Medicaid who remain under normal fee-for-service reimbursement coverage and who are users of OPDUR pharmacies. Most Medicaid recipients who are permanently disabled or elderly will be Medicare entitled – dually eligible. Since Medicare is the primary payer of non-pharmacy services for these population, Medicare claims data is the source of health outcomes information and Medicaid is the sources of pharmacy use information.

---

[15] The SAS System's CATMOD procedure was used with the NOGLS option specified.

### *Definitions of Baseline and Follow-up Periods*

Pharmacies enrolled in the demonstration in three waves (February, April, and September of 1994). Although other approaches would be possible, this study assumes that a patient's exposure to the demonstration (and consequently the patient's follow-up period) began when a patient first filled a prescription at a pharmacy participating in the demonstration. In the majority of cases, the follow-up period then continued until the evaluation data collection ended in 1996. In certain cases, if the patient lost eligibility for Medicaid, entered a nursing home, or entered an HMO, the follow-up period ended sooner because the outcomes data captured in the claims would be incomplete after that event. A patient's death was also deemed to terminate the follow-up period.

In a similar way, patients' baseline periods differed in length. The baseline period was determined to end the day before the patient's first contact with a demonstration pharmacy and to reach back twelve months or until the patient first became eligible for became eligible for the Medicaid program or left an HMO. The restrictions are imposed in order to ensure that the medical record is complete during the baseline period. Clinical outcomes that occurred on the day the patient first went to a demonstration pharmacy are excluded from both the baseline and the follow-up periods. In such cases, it is difficult to determine whether the clinical issue precipitated the prescription or the prescription led to the clinical issue without a medical records review.

### *Dependent Variables*

Based on the research literature, the GI events analyzed in previous studies relate to a hospitalization event with a upper GI ulcer or hemorrhage diagnosis [ Carson et.al. (1987)]. Several studies narrowed the outcome definition and focused on death subsequent to a GI hospitalization [Griffin et. al. (1988); Guess et. al. (1988)]. Because of the increase in ambulatory surgical treatments, the outpatient or inpatient occurrence of a GI endoscopy with a GI ulcer or hemorrhage diagnosis was also identified as a analyzable outcome [Duggan et. al. (1986)]. The epidemiological evidence focused on measuring severe effects associated with NSAID exposure the requirements of the analysis of the demonstration program argue for a broader review of the impact of a pharmacy intervention on both the incidence and prevalence of GI events and symptomology. The review of follow-up experience subsequent to the first use of a demonstration pharmacy differs in scope and nature from testing for severe GI events after an NSAID exposure. Improvements in pharmacy practice related to better supervision of NSAID use would be expected to be manifested in the reduction of severe GI outcomes, but also, more broadly, in the alleviation in GI distress as reflected in a reduction of reported GI symptoms as seen in outpatient diagnoses. The analyses reviewed outcomes that spanned the range of identifications of GI outcomes using the information available in claims data. The uncertainties in accurate diagnostic coding and the possible confounding relationship between demonstration pharmacy use and physician visit patterns for minor symptoms may serve to confound the most general measures.

**Exhibit 5.8.1**

**HCFA DUR Demonstration Evaluation**
**GI Outcome Markers**

| Outcome Type | Data Source | ICD9/CPT |
|---|---|---|
| Primary inpatient hospital diagnosis for ulcer or gastrointestinal hemorrhage followed by death within 30 days | Medicaid inpatient hospital primary diagnosis , Medicare inpatient hospital primary diagnoses and date of death from Medicaid eligibility records or Medicare enrollment records | ICD9 531-534, 578 |
| Primary inpatient hospital diagnosis for ulcer or gastrointestinal hemorrhage | Medicaid inpatient hospital primary diagnosis , Medicare inpatient hospital primary diagnoses | ICD9 531-534, 578 |
| Any inpatient hospital diagnosis for ulcer or gastrointestinal hemorrhage | Medicaid inpatient hospital diagnosis , Medicare inpatient hospital diagnoses | ICD9 531-534, 578 |
| Any hospital or physician diagnosis for ulcer or gastrointestinal hemorrhage with a GI endoscopy | Medicaid or Medicare hospital, outpatient or physician diagnosis and procedure codes | ICD9 531-534, 578<br>CPT 432 |
| Any hospital or outpatient diagnosis for ulcer or gastrointestinal hemorrhage | Medicaid or Medicare hospital, outpatient or physician diagnosis | ICD9 531-534, 578 |

*Sources:* Review of literature of GI events and NSAID exposure

Outcome indicators were constructed using Medicaid and Medicare diagnosis. Each outcome was associated with a series of ICD-9 codes. If a patient's record showed evidence of one of those ICD-9s as either the primary or secondary diagnosis for an inpatient stay or for a physician encounter (paid by Medicaid or Part A or Part B of Medicare) during the follow-up period, then the patient was determined to have had the outcome.

### *Independent Variables – Clinical Risk Factors*

Each model included a basic set of covariates that captured exposure to the intervention as well as general predisposing and need characteristics. Additionally, covariates were added to each model to control for clinical risk factors that are specific to the outcomes. The main set of independent variables, as well as the outcome specific independent variables, are described in the following sections.

Based on a review of the literature a broad selection of factors related to GI outcomes were identified. The major factors were related to a prior history of GI disease, disease that increase the risk of GI outcomes and conditions whose treatment is linked to increased risk of GI outcomes. Because the source of the covariate information was administrative claims data, diagnoses and pharmaceutical treatments related to the diagnoses were both used for the assessment of individual status.

**Exhibit 5.8.2**

**HCFA DUR Demonstration Evaluation**
**Clinical Covariates**[16]

| Group Name | Description |
|---|---|
| GI1 | physician or hospital; diagnosis for upper GI ulcer |
| | physician or hospital diagnosis for GI hemorrhage |
| | H2-antagonist and others |
| | antacids |
| | Misoprostol |
| GI2 | non-infective enteritis |
| | colitis |
| | diverticulitus |
| GI3 | disease of esophagus |
| | gastritis |
| | disorders of stomach |
| | gastroduodenal disorders |
| | other peritoneal disorders |
| CRD1 | hypertension |
| | anti-hypertensives |
| | beta-blockers |
| | calcium channel blockers |
| CRD2 | cardiotonics |
| | anti-anginals |
| | anti-arrhytmics |
| DB | diabetes |
| | insulin, oral hypoglycemics |
| CR | chronic renal failure |
| CP | chronic obstructive pulmonary disease/asthma |
| CS | use of corticosteroids |
| AN | anemia |
| CD | coagulation disorders |
| | anti-coagulant |
| AL | hepatic disease |
| | alcoholism |
| NS | NSAID |

*Sources:* Review of literature of GI events and NSAID exposure

Although the demonstration was designed as a random experiment, randomization occurred at the level of the pharmacy cluster while clinical benefits accrued at the level of the patient.[17] This posed several methodological challenges. First, individuals had the potential to go to multiple pharmacies and, potentially, to go to a mix of intervention, control, and non-demonstration pharmacies. The study adopted a simple approach. Individuals were considered members of the intervention group if they ever had a Medicaid-paid drug claim at an intervention pharmacy and never had such a claim at a control pharmacy. They were considered members of the control group if they ever had a Medicaid-paid drug claim at a control pharmacy and never had such a claim at an intervention pharmacy. Individuals who went to both intervention and control pharmacies were excluded from the study.[18]

---

[16] Diagnosis codes are specified to the third digit only. Reviews of coding of the ICD9 manual and observed patterns suggest that broadly defined groups are appropriate for the conditions listed for covariates. The actual selection of drug covariates was based on NDC lists prepared by Pennsylvania State University research team.

[17] Pharmacy clusters, built around common prescribers, were created and randomized to minimize the influence that I intervention pharmacies might exert through common prescribers on control patients. In Iowa, there were n pharmacy clusters ranging in size from n person to npeople.

The basic design of the study was to compare treatment and control patients within a regression framework that controls for demographic, drug, and disease factors. For all outcomes, regressions included controls for the patient's race, sex, and age. These factors affect health outcomes both directly and as proxies for behavior; they may also affect the predisposition to use care. These variables, which were taken from the enrollment records of the Medicaid program, were included in regressions as a series of indicator variables.

### *Independent Variables – Length of Baseline and Follow up Periods*

A patient with a longer follow-up period would be more likely to experience any clinical outcome than one with a shorter follow-up period, thus the length of the follow-up period was made an additional control in models. While the length of the baseline period does not affect patients' probabilities of clinical problems in and of itself, expected health spending, number of drugs, and risk factors will tend to be slightly under measured for patients with short baselines relative to patients with full baselines. Thus, controlling for measured covariates, the probability of clinical outcomes might be higher for patients with short baselines; consequently, the length of the baseline period was included as an additional control. The lengths of the follow-up and baseline period were entered in models as the natural log of the number of months in each period.

The analyses of GI outcomes in much of the medical literature are based on short duration time correlations between NSAID exposure and GI events. It has been suggested that the impact of NSAIDs is greatest in short term, high dose regimens. The hypothesis in this analysis is that an improvement in pharmacy practice will lead to improved health status. A potentially long duration study period is justified since the relationship between the entry of a study subject into the OPDUR coverage and the health benefits are not time defined in the same way as the clinical relationship between an NSAID exposure and a GI outcome. Patient counseling of GI distress management, drug therapy dose modification, therapy substitutions and prophylactic treatment may occur over the follow-up and result in long term benefits observed as either decreases in severe outcomes or reductions in the symptoms of iatrogenic GI effects.

### *Independent Variables – Pharmacy Wave*

Pharmacies entered the computerized utilization review demonstration in two waves. The first wave of 160 pharmacies enrolled in February 1994. A second wave (N = 33) of pharmacies entered in September 1994. The first wave was used as the reference population in the model.

The basic set of independent variables that were used in every model (as described above) are summarized below.

---

[18] For this purpose, the day that the patient first went to the demonstration pharmacy was included in the follow-up period.

**Exhibit 5.8.3**

**HCFA DUR Demonstration Evaluation**
**Basic Set of Covariates Included in All Models**

race (in categories: caucasian, non-caucasian)
sex
age (in categories: 45-64, 65-74, 75-84, 85+)
length of baseline period
length of follow-up period
wave contact demonstration pharmacy (first wave, other)

*Risk Groups*

A variety of demonstration sub-populations were identified from the relevant medical literature that were at greater risk for NSAID induced GI events. These groups are enumerated below in Exhibit 5.8.4. The regression analyses were separately applied to the general population and to the risk groups. The risk groups include populations that have been identified in the literature as sensitive to NSAID induced GI outcomes. The lower age cutoff used for all groups was 45 years. Greater age is generally associated with higher risk of GI outcomes, studies have been done based on cutoffs of 47, 60 and 75. The other rationale for an age limit is that the Medicaid program covers very specific populations, the younger groups are primarily women with young children who are not considered at risk for NSAID induced GI effects. The older populations, after the exclusion for nursing home residency, are a mixture of the community dwelling low income disabled and elderly.

**Exhibit 5.8.4**

**HCFA DUR Demonstration Evaluation**
**Study Populations**

| Study Population | Description |
|---|---|
| Main Population | Eligible for either Medicare and Medicaid, at least 45 years old in baseline period, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies. |
| NSAID User | Eligible for either Medicare and Medicaid, at least 45 years old in baseline period, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled an NSAID prescription in baseline period. |
| Any GI Diagnosis or Related Prescription | Eligible for either Medicare and Medicaid, at least 45 years old in baseline period, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, physician or hospital diagnosis of GI ulcer or hemorrhage or a prescription filled for a H2-Blocker, antacid or Misoprostol in the baseline period. |

*Sources:* Iowa Medicaid and Medicare claims and enrollment records

## Results

The users of demonstration of pharmacies were assigned to study groups depending on the OPDUR status of their pharmacy. The OPDUR pharmacies were assigned to the treatment and control groups based on a cluster randomization methodology that was designed to minimize the overlaps in the physician-prescriber community potentially affected by demonstration influences. While the randomization was pharmacy based the outcomes analyses focused on individual patients. Below are tables of the population characteristics of the study groups in the baseline period for key covariates and the results of demographic and disease adjusted analyses.

**Exhibit 5.8.5**

**HCFA DUR Demonstration Evaluation**
**Demonstration Descriptive Statistics**

| Selected Baseline Statistics | Experimental | Control | T-Test Probability Prob>ITI Unequal |
|---|---|---|---|
| Study Subjects | 6,386 | 6,778 | |
| Person-Months | 61,852 | 66,263 | |
| Mean Age | 63.9 | 64.4 | 0.0272 |
| Mean Gender (0=M/1=F) | 0.72 | 0.73 | 0.0279 |
| Mean Baseline Months | 9.7 | 9.8 | 0.0001 |
| Mean Follow-up Months | 17.6 | 17.1 | 0.0001 |
| Primary Diagnosis of Gastric Ulcer | 0.2% | 0.3% | 0.0656 |
| Primary Diagnosis of GI Hemorrhage | 0.2% | 0.2% | 0.5002 |
| Diagnosis of Gastric Ulcer | 2.8% | 2.9% | 0.8072 |
| Diagnosis of GI Hemorrhage | 2.5% | 2.9% | 0.1519 |
| GI endoscopy Procedure | 3.2% | 3.0% | 0.3622 |
| Arthritis Diagnosis | 29.9% | 30.9% | 0.0101 |
| Diabetes Diagnosis | 15.2% | 16.5% | 0.0337 |
| COPD/Asthma Diagnosis | 15.0% | 15.1% | 0.9621 |
| CRF Diagnosis | 1.2% | 1.4% | 0.2567 |
| Alcoholism Diagnosis | 4.6% | 4.0% | 0.1044 |
| Hypertension Diagnosis | 24.2% | 28.0% | 0.0006 |
| NSAID Rx | 32.1% | 32.8% | 0.4482 |
| H2-Blocker Rx | 24.7% | 25.2% | 0.4680 |
| Misoprostol Rx | 0.8% | 0.7% | 0.2718 |
| Antacid Rx | 0.0% | 0.0% | 0.3173 |
| Corticosteroid Rx | 9.7% | 10.1% | 0.4624 |
| Anti-Coagulant Rx | 4.8% | 5.0% | 0.7213 |

*Sources:* Iowa Medicaid and Medicare Claims and Enrollment Data 1993-1996

### *Baseline Statistics*

The use of NSAIDs in the baseline period coupled with a GI diagnosis or another drug therapy that can increase the risk of GI bleeding may be indicative of pharmacy practices that can be improved. The statistics in Exhibit 5.8.6 are measures of problematic sub-populations as distributed by their affiliation with OPDUR pharmacy types at the time of demonstration induction. Similarly, the rate of PSU/UMPCPS screen violations for NSAID therapies can be used to identify high dosage or duplicate regimens. The statistics are measures of concurrent prescriptions filled in the same month and diagnoses reported in the same month as a NSAID prescription fill or in the subsequent month.

**Exhibit 5.8.6**

**HCFA DUR Demonstration Evaluation**
**Descriptive Statistics**

| NSAID-GI Baseline Statistics | Experimental (N=6,386) | Control (N=6,778) | T-Test Probability Prob>ITI Unequal |
|---|---|---|---|
| NSAID - Ulcer or Hemorrhage | 1.0% | 1.4% | 0.0351 |
| NSAID - H2-Blocker Rx | 9.3% | 9.6% | 0.4951 |
| NSAID - Antacid Rx | 0% | 0% | - |
| NSAID - Misoprostol Rx | 0.7% | 0.5% | 0.1239 |
| NSAID - Anti-Coagulant Rx | 1.1% | 1.2% | 0.5881 |
| NSAID - Corticosteroid Rx | 2.7% | 2.8% | 0.7867 |
| NSAID PSU Dosage Violation | 10.9% | 10.7% | 0.8159 |

*Sources:* Iowa Medicaid and Medicare Claims and Enrollment Data 1993-1996

The populations assigned to the two pharmacy types are generally similar in personal demographics and in the length of baseline and follow-up periods, although the size of the study groups leads to statistical significance for small differences. Likewise there are small but significant differences in the observed rate of several of the chronic disease diagnosis categories. GI ulcer and hemorrhage diagnoses and treatments do not vary significantly between the two study groups. The one significant deviation of interest is the frequency of concurrence between an NSAID prescription fill and the reporting of an ulcer or GI hemorrhage diagnosis is elevated in the control group during the baseline period.

### *Regression Results*

In total five models were prepared for the main population and two risk groups, resulting in ten regression analyses. The parameter estimates for all variables used in the logistic regression of the main population are included in Appendix F. The demographic and clinical covariates generally showed the anticipated relationship to GI outcomes. Three factors that consistently showed a non-significant relationship to any GI outcomes were withdrawn from the models, they were: 1) exposure to an anti-coagulant medication or an coagulation disorder diagnosis in the baseline period; 2) the drugs grouped under CRD2; 3) the diagnoses grouped under GI3.

Below are presented the adjusted odds ratios derived from the covariance matrix for each model in the general population and two of the risk groups. The main covariate influences were observation time, longer follow-up leads to greater risk, increased age is correlated with greater risk of GI outcomes, gender, female reduces risk and a baseline GI diagnosis is highly associated with the occurrence of events in the study period. Baseline diagnoses of renal failure, anemia, alcoholism were all strongly correlated with GI bleed hospitalizations.

**Exhibit 5.8.7**

**HCFA DUR Demonstration Evaluation**
**Logistic Regression Results for OPDUR Status**

| **General Population Age>45,** | **OR** | **95% CI** |
|---|---|---|
| Hospitalization with a primary diagnosis of ulcer or GI | 1.94 | 0.54-7.00 |
| Hospitalization with a primary diagnosis of ulcer or GI | 0.76 | 0.53-1.09 |
| Inpatient diagnosis of ulcer or GI hemorrhage | 0.90 | 0.71-1.13 |
| Inpatient or outpatient diagnosis of ulcer or GI hemorrhage and | 1.08 | 0.86-1.38 |
| Inpatient or outpatient diagnosis of ulcer or GI hemorrhage | 0.83 | 0.73-0.94 |
| **Population Age>45, NSAID User in Baseline Period,** | **OR** | **95% CI** |
| Hospitalization with a primary diagnosis of ulcer or GI | - | - |
| Hospitalization with a primary diagnosis of ulcer or GI | 1.00 | 0.56-1.77 |
| Inpatient diagnosis of ulcer or GI hemorrhage | 1.13 | 0.78-1.62 |
| Inpatient or outpatient diagnosis of ulcer or GI hemorrhage and | 1.14 | 0.79-1.65 |
| Inpatient or outpatient diagnosis of ulcer or GI hemorrhage | 0.80 | 0.66-0.97 |
| **Population Age>45, Any GI Ulcer/Hemorrhage Diagnosis,** | **OR** | **95% CI** |
| Hospitalization with a primary diagnosis of ulcer or GI | 2.22 | 0.19-25.58 |
| Hospitalization with a primary diagnosis of ulcer or GI | 0.78 | 0.45-1.33 |
| Inpatient diagnosis of ulcer or GI hemorrhage | 0.95 | 0.69-1.30 |
| Inpatient or outpatient diagnosis of ulcer or GI hemorrhage and | 1.04 | 0.75-1.45 |
| Inpatient or outpatient diagnosis of ulcer or GI hemorrhage | 0.79 | 0.66-0.94 |

*Sources:* Iowa Medicaid and Medicare Claims and Enrollment Data 1993-1996

The adjusted odds ratios derived from the regression parameter estimates show no significant result for OPDUR pharmacies in the A group in decreasing the risk of severe GI outcomes. The statistical power for determining the effect of computerized utilization review on death subsequent to a hospitalization for GI ulcer or hemorrhage is insufficient in the study population for any definitive measurement of risk. Hospitalizations with a primary diagnosis of GI bleed were not significantly decreased. Reports of GI bleed diagnoses during inpatient stays had non-significant decreases in risk in all models. Endoscopy procedures were not significantly reduced.

The only consistent positive result was an apparent significant decrease in the reporting of a GI bleed diagnosis in all settings in the general population and in the two risk groups. The mechanism for a decrease in risk for general GI symptomology could follow two tracks. The first is that changes is therapeutic prescribing regimens resulted in lower rates of adverse effects. The second track is that the symptomology remained the same, however pharmacy base computerized utilization review reduced the need for physician visits, perhaps by increasing consultations between physicians and pharmacists. In either case the analysis suggests a possible benefit in a decrease in GI problems correlated to OPDUR pharmacy status.

## Discussion

There are external and internal factors that could explain the decrease in GI diagnoses in all settings. An important explanatory element would be the connection between modifications in patient drug therapies related to the decrease in risk. Because many of the important drugs are over the counter some of the most possible modifications, e.g. elimination of aspirin, use of antacids and others are essentially unobservable. Additionally diagnostic information is sensitive to physician reporting practices and since physicians were not randomized there may be biases in the distribution of physician specialties and practice patterns in the experimental groups that would cause differential reporting of GI diagnoses.

Below is presented a analysis of several therapy modalities which are considered to be problematic - NSAID with concurrent corticosteroid, NSAID with an anti-coagulant, PSU/UMPCPS NSAID screener violation - as well as statistics relating to the populations using NSAIDs, H2-blockers and Misoprostol in the follow-up period. The statistics are measures of concurrent prescription dispense events in the same month and diagnoses in the same month or in the subsequent month. The measurements are unadjusted for demographic and disease covariates, however they are illustrative of the extent to which the demonstration affected specific therapeutic combinations. The measurements are based on binary variables of the presence or absence of drug use in the follow-up period.

**Exhibit 5.8.8**

**HCFA DUR Demonstration Evaluation**
**Descriptive Statistics**

| NSAID-GI Follow-up Statistics | Experimental (N=6,386) | Control (N=6,778) | T-Test Probability Prob>ITI Unequal |
|---|---|---|---|
| NSAID - Ulcer or Hemorrhage | 2.9% | 3.2% | 0.4190 |
| NSAID - H2-Blocker Rx | 14.2% | 14.3% | 0.9780 |
| NSAID - Misoprostol Rx | 0.9% | 0.9% | 0.8294 |
| NSAID - Anti-Coagulant Rx | 2.1% | 2.1% | 0.8976 |
| NSAID - Corticosteroid Rx | 5.3% | 4.8% | 0.1435 |
| NSAID PSU Dosage Violation | 36.4% | 36.8% | 0.6964 |

*Sources:* Iowa Medicaid and Medicare Claims and Enrollment Data 1993-1996

A logistic regression analysis using the same covariates as for the analysis of GI effects, shows no significant elevations in Misoprostol or H2-blocker therapy concurrent with an NSAID prescription in the treatment pharmacies during the follow-up period. Drug combinations that are potentially problematic, NSAID – corticosteroids, NSAID – anti-coagulants, were not significantly different in the treatment and control groups. There is a large increase, relative to the baseline statistics, of NSAID dosage screen violations but they occur equally in the study groups.

The randomization of patients based on minimization of overlap in prescriber communities could have resulted in biases in the reporting and treatment of GI symptoms. The use of diagnoses from all settings as a method for measuring disease prevalence can be problematic if there are strong community variations in the coding and reporting of secondary symptoms that result in observation bias. Outcomes which are not well defined events are especially sensitive to these factors.

Similarly, community practices of prescribers and pharmacies can be vary during the baseline and evolve differentially over the follow-up period, independent of the effect of the OPDUR program. Changes in community standards that do not equally affect the study populations cannot easily be adjusted for, however pre-existing differences can be measured in a review of the baseline data.

An analysis corrected for demographic characteristics of differences in the two study groups during the baseline and follow-up periods is presented below. The statistic is listed as the odds ratio for GI events prior to the demonstration by study group adjusted by demographic and observation time factors only. The question that is addressed is whether the observed effect in the treatment population pre-dates the start of the demonstration.

**Exhibit 5.8.9**

**HCFA DUR Demonstration Evaluation**
**Logistic Regression Results for OPDUR Status**

| **General Population Age>45** | **OR** | **95% CI** |
|---|---|---|
| Hospitalization with a primary diagnosis of ulcer | 0.68 | 0.40-1.16 |
| Inpatient diagnosis of ulcer or GI hemorrhage | 0.92 | 0.68-1.25 |
| Inpatient or outpatient diagnosis of ulcer or GI | 1.00 | 0.72-1.41 |
| Inpatient or outpatient diagnosis of ulcer or GI | 0.94 | 0.80-1.10 |
| **General Population Age>45** | **OR** | **95% CI** |
| Hospitalization with a primary diagnosis of ulcer | 0.75 | 0.53-1.07 |
| Inpatient diagnosis of ulcer or GI hemorrhage | 0.88 | 0.70-1.10 |
| Inpatient or outpatient diagnosis of ulcer or GI | 1.05 | 0.83-1.32 |
| Inpatient or outpatient diagnosis of ulcer or GI | 0.82 | 0.72-0.92 |

*Sources:* Iowa Medicaid and Medicare Claims and Enrollment Data 1993-1996

The baseline analysis does not demonstrate a significant pre-existing difference in the reporting of GI bleed diagnoses between the treatment and control pharmacies. This observation tends to support the conclusion of a positive treatment effect.

## Conclusion

An OPDUR demonstration impact on severe GI outcomes cannot be asserted, there were no statistically significant differences between the experimental groups in inpatient primary diagnoses for

GI bleeding in the study period. Additionally the risk of a severe GI episode does not change between the baseline and follow-up time windows. Similarly reports of GI ulcer and hemorrhages as observed in hospital diagnoses overall are not affected by OPDUR pharmacy status. The one measurement of difference appears in the observation of GI bleed diagnoses collected from hospitals and physicians in all settings. The protective effect indicates a 20 percent reduction in the probability of a report of a GI symptom. A review of possible intermediation mechanisms — drug prophylaxis, reduction in problematic therapeutic combinations — does not suggest a reason for the reduction in risk. A decrease in reports of GI bleed diagnoses could be sensitive to pharmacist linked service substitutions as opposed to a true decline in symptomology. Whether the decrease is in actual GI disease or in physician and hospital services where GI diagnoses are reported, the effect is of interest.

## 5.9 The Impact of the OPDUR Demonstration Project on Adverse Clinical Outcomes Related to the Use of Antidepressants

Inappropriate prescribing can lead to costly, serious adverse health outcomes. Prospective drug use review is intended to prevent these adverse health outcomes by alerting pharmacists to drug therapy problems at the time the prescription is filled. Unlike retrospective drug use review (RDUR), which is conducted after the prescription has been dispensed, prospective drug use review has the potential to resolve drug therapy problems before they result in illness or even death.[Chrischilles et. al. (1997)].

As part of the Omnibus Budget Reconciliation Act of 1990 (OBRA '90), the Health Care Financing Administration funded a demonstration of on-line prospective drug utilization review (OPDUR) in Iowa. OPDUR is a form of prospective drug use review that relies on computerized algorithms to screen prescriptions for problems. Usually within seconds, and at the point of service, an OPDUR system alerts the pharmacist to drug therapy problems that have been identified in the prescription being screened.

The Iowa OPDUR system utilizes information about prescriptions filled at any pharmacy in the state that accepts Medicaid; this feature of the system allows the pharmacist to make more informed decisions that would be possible with a standard pharmacy computer system. Furthermore, the Iowa OPDUR algorithms are based on the University of Maryland and Philadelphia College of Pharmacy and Sciences (UM/PCPS) criteria, which constitute the most refined set of criteria currently available. It was hypothesized therefore that the OPDUR system in Iowa would be effective in preventing drug therapy problems which have the potential to result in costly and potentially serious adverse health outcomes.

The Iowa OPDUR system was in operation between June 1993 and June 1997. Pharmacies entered the demonstration in two waves, and were randomly assigned to treatment and control groups in clusters. This study assessed the effects of the Iowa OPDUR demonstration on a set of adverse clinical outcomes that have been linked to inappropriate prescribing of antidepressant drugs. Specifically, it explored the impacts on cardiac conditions, orthostatic hypotension (sometimes marked by hip fractures), and a cluster of anticholinergic symptoms; the clinical literature indicates that all of these conditions are associated with the use of antidepressant drugs. The population of interest in this study was elderly Iowa residents who were eligible for Medicare and Medicaid, and who filled a prescription at a demonstration pharmacy at least once.

The results of the analysis suggested that the Iowa OPDUR system only had a marginal impact on the anticholinergic outcome and did *not* have a measurable effect on the other adverse outcomes related to antidepressant drug use. Holding baseline characteristics constant, individuals who visited pharmacies in the treatment group were just as likely as control group patients to experience cardiac effects and hypotension symptoms. These findings suggest that while the Iowa OPDUR system may have led to benefits in other areas, the intervention was not effective in improving antidepressant drug prescribing quality to the extent that major, adverse health outcomes were prevented.

## Adverse Clinical Outcomes Related to Antidepressants

Numerous studies support the conclusion that the use of tricyclic antidepressants (TCAs) at high doses is associated with serious cardiac effects. This class of drugs can affect the conduction system, increase the heart rate, and exert an effect on heart rhythm. Some researchers have attempted to identify whether or not these effects are observed at therapeutic doses as well as high doses. The research literature also links TCAs with orthostatic hypotension and anticholinergic effects.

All the outcomes under study are associated with tricyclic antidepressants, and not with other classes of antidepressants such as MAO inhibitors or targeted antidepressants, such as SSRIs. Initially, the outcomes were identified based on associations with the PSU screening criteria; these criteria emphasize TCAs. Also, because TCAs were less costly than SSRIs during the study period, they probably were the dominant class of antidepressant among the Medicaid population

### *Cardiac Effects*

***Conduction Disorders.*** TCAs can prolong cardiac conduction, but this event will almost never produce symptomatic heart disease at therapeutic levels for patients without pre-existing conduction disorders [Jackson et al.(1987)]. By slowing intraventricular conduction, TCAs can lead to complete heart block or ventricular reentry arrhythmias in cases of overdose or among patients with pre-existing cardiac disease [Glassman and Preud'homme (1993)]. At overdose levels, conduction delays and heart block can be significant, even fatal problems. For patients who already have conduction disorders, the conduction delays produced by TCA use can lead to additional cardiac problems, including fatal heart block [Roose et al.(1987)]. Among patients taking TCAs, those with pre-existing bundle branch blocks are thought to be at greatest risk of developing high degree atrioventricular conduction delays [Jackson et al. (1987)]. Jackson et al. (1987) state: "Apart from orthostatic hypotension, the major risk of the tricyclic antidepressants is tricyclic-induced atrioventricular block in patients with pre-existing conduction disease."

***Heart Rhythm Effects.*** TCAS have a class IA antiarrhythmic action, and several studies have indicated that drugs with class IA and/or IC antiarrhythmic action increase the risk of mortality among patients with arrhythmias [Glassman (1993); Morganroth et al. (1991)]. Glassman (1993) further states that there has been recent evidence from additional studies that antiarrhythmic drugs may pose a mortality risk when used to suppress atrial fibrillation, as well as when used to suppress ventricular arrhythmias after MI.

The mechanism by which class I antiarrhythmic drugs produce increased mortality is not known [Glassman (1993)]. However, it appears that a likely candidate is an interaction between a class IA antiarrhythmic drug and an ischemic myocardium. Glassman and Roose (1994) present the current understanding of this relationship, "Class I antiarrhythmic drugs are strongly antiarrhythmic under aerobic conditions. However, during periods of oxygen deprivation, such as angina or myocardial infarction, these drugs appear to become paradoxically proarrhythmic."

The proarrhythmic effect can be expected to be seen in approximately 10 percent of patients with pre-existing arrhythmias who are given class I antiarrhythmics [Glassman and Roose (1994)]. Glassman and Roose go on to warn, however, that it is not only patients with pre-existing arrhythmia who have ischemic hearts. Avoiding TCA use in patients with arrhythmia will not protect patients with asymptomatic arrhythmias or ischemic heart disease but no arrhythmia. They further say that from a point of view of safety, it would be best to avoid a TCA "in any patient who is at serious risk for ischemic heart disease."

### *Orthostatic Hypotension*

Orthostatic hypotension is thought to be the most common of the serious side effects of TCA use at therapeutic levels. It is almost universally reported in articles written about the side effects associated with TCAs. Roose et al. (1987) state that orthostatic hypotension occurs in 5 percent to 10 percent of patients taking TCAs. In depressed patients who have impaired left ventricular function and are treated with imipramine the rate of orthostatic hypotension can be much higher [Glassman et al. (1983); Chutka (1990)].

Consequences of hypotension include falls and fractures; as Glassman (1984) points out, major drops in blood pressure occur at all ages, but the risk of serious sequelae increases considerably with age. Two studies by Ray et al. (1987 and 1991) examine the association of hip fracture (one of the major serious sequelae in the elderly) and TCA use. Ray et al. (1987) found that the relative risk of hip fracture for TCAs was 1.9 among the Medicaid population (95 percent Confidence Interval is 1.3-2.8). Ray et al. (1991) found the relative risk for current TCA use was 1.6 among an elderly Canadian population (95 percent Confidence Interval is 1.3-1.9). The study also indicated that the risk of hip fracture was greatest in the first 90 days of therapy, though there was increased risk for longer durations of use as well.

### *Anticholinergic Effects*

TCAs also produce anticholinergic effects, the most common of which are dry mouth, blurred vision, constipation, tachycardia, and urinary hesitancy/retention [Coccaro and Siever (1985)]. A number of studies and reviews discuss the occurrence of anticholinergic effects. Perhaps the most recent is an article by Simon et al. (1996) in which fluoxetine (an SSRI) is compared with imipramine (a tertiary amine TCA) and desipramine (a secondary amine TCA) using clinical, functional, and economic outcomes. Constipation, GI upset, headache, stimulation and dizziness occurred more frequently among TCA patients.

## Evaluation Approach

The evaluation approach was to explore the relationship between the Iowa OPDUR system and the rate of adverse events related to use of antidepressant drugs. It was hypothesized that pharmacists using the OPDUR system would be more likely than pharmacists not using the OPDUR system to detect prescription problems related to antidepressant drug use and, consequently, to have an opportunity to prevent the occurrence of adverse clinical outcomes. Patients who visited pharmacies where prescriptions were screened by OPDUR, therefore, should be less likely than control group patients to experience illnesses related to inappropriate prescribing of antidepressant drugs.

The basic design was to compare rates of clinical outcomes among patients who went to intervention pharmacies to those among patients who went to control pharmacies. Multiple regression techniques were used in order to control for demographic, disease, and drug factors predisposing a given patient to the outcome under study. All control variables were measured during a baseline period before the demonstration began so as to be independent of any demonstration effects. A version of the Huber/White variance estimator that incorporates cluster-based sampling was used to estimate standard errors [Rogers (1993)].

### *Data*

The data used in this study was extracted from Medicaid drug claims and Medicare health services utilization claims for patients in the state of Iowa in 1993 (to provide a baseline) and 1994-1996. (OPDUR was in operation in Iowa between June 1993 and June 1997) Other records and data streams generated by the demonstration were also incorporated into the data set. Data development is described in detail in Chapter 5.7.

### *Study Populations*

As displayed below in Exhibit 5.9.1, the data was used to create four study populations: (1) a main study population; (2) a sub-sample of the main study population consisting of those who had been using antidepressants during the baseline period; (3) a sub-sample of the main population consisting of those who had been using antidepressants during the baseline or the follow-up period; and (4) a sub-sample of the main study population whose prescriptions triggered screen failures related to the use of antidepressant drugs. Although all models were implemented using these four samples, not all results are presented and discussed in this chapter. Instead, this chapter focuses on interpreting and summarizing the results from these models; selected models estimated using the main population are presented in full in the appendix to this chapter.

The main study population consisted of elderly patients (65 and over on January 1, 1993) who were eligible for both the Medicaid and Medicare programs and who went to either intervention or control pharmacies during the intervention. The elderly are particularly at risk for both the cardiac and hip fracture outcomes. Moreover, they are generally in more delicate health and taking more drugs than younger patients; the elderly thus are hypothesized to be more likely than people of other age groups to benefit from the intervention. The requirement of dual eligibility ensures that the analyses have complete and consistent records of drug and health services utilization for all patients in the data set.

**Exhibit 5.9.1**

## HCFA DUR Demonstration Evaluation Antidepressant Study Populations

| | |
|---|---|
| Main Population<br>(N = 4220) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies |
| Baseline or Follow-up Antidepressant Users<br>(N = 1042) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled at least one prescription for an antidepressant during the baseline period or during the follow-up period |
| Baseline Antidepressant Users<br>(N = 690) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled at least one prescription for an antidepressant during the baseline period |
| Triggered an Antidepressant Drug Screen Failure (N = 331) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, triggered at least one screen failure related to antidepressant drug use. |

*Sources:* IA Medicaid claims

The study population was further limited to individuals who used antidepressants during the baseline period, and individuals who used antidepressants during the baseline or follow-up periods. Narrowing the population in this way was intended to improve the specificity of the analysis.[19] The population of all antidepressant users (rather than TCA users alone) was used to retain patients for whom the intervention effect involved the substitution of another antidepressant for a TCA in order to avoid the clinical risks associated with TCAs. Furthermore, the analysis considered both a sub-sample of baseline antidepressant users as well as a sub-sample of baseline or follow-up antidepressant users because of concerns that the intervention might affect antidepressant use in the follow-up period. In this case, estimates based on the baseline/follow-up antidepressant sub-sample might be biased.[20]

The models also were estimated using a sub-sample that consisted of those patients whose prescriptions had triggered an antidepressant drug screen failure. In the treatment group, the pharmacist received the screen failure information on his/her computer and presumably read and reacted to the information.[21] In the control group, screen failure information was *not* transmitted to the pharmacist. Instead, in the control group, the screen failure information was recorded for research purposes only. In the sub-population of persons who triggered screen failures, therefore, this analysis compares the rate of adverse outcomes between individuals whose pharmacists were alerted via OPDUR to an antidepressant drug prescription problem and individuals whose pharmacists were not alerted via OPDUR to an antidepressant drug prescription problem. The screen-triggered sample is the most targeted in the sense that every individual in the sample had an antidepressant drug prescribing problem; compared to the other samples analyzed, this group was at the highest risk of adverse effects. Nevertheless, the sample size is reduced considerably by excluding those who did not trigger a screen failure.

---

[19] Nevertheless, analyses based on this more narrow population would miss any intervention effect that involved patients avoiding the use of antidepressants altogether.

[20] If the Iowa OPDUR system affected antidepressant use in the follow-up period and the population of interest consists of antidepressant users in the baseline and follow-up periods, a "selection problem" would exist. In this case, it is possible that those who benefitted from the intervention would be more likely than other to be excluded from the sample.

### *Definitions of Baseline and Follow-up Periods*

Pharmacies enrolled in the Iowa OPDUR demonstration in two waves (June 1994 and December 1994). Although other approaches would be possible, this study assumes that a patient's exposure to the demonstration (and consequently the patient's follow-up period) began when a patient first filled a prescription at a pharmacy participating in the demonstration. In the majority of cases, the follow-up period then continued until the last month of 1996.[22] In certain cases, if the patient lost eligibility for Medicare or Medicaid, entered a nursing home, or entered an HMO, the follow-up period ended sooner because the outcomes data captured in the claims would be incomplete after that event.[23] A patient's death was also deemed to terminate the follow-up period.

In a similar way, patients' baseline periods differed in length. The baseline period was determined to end the day before the patient's first contact with a demonstration pharmacy and to reach back twelve months or until the patient first became eligible for the Medicare program, became eligible for the Medicaid program, left an HMO, or left a nursing home and became resident in the community. Again, the latter three restrictions are imposed in order to ensure that the medical record is complete during the baseline period.

Note that clinical outcomes that occurred on the day the patient first went to a demonstration pharmacy are excluded from both the baseline and the follow-up periods. In such cases, it is impossible to know whether the clinical issue precipitated the prescription or the prescription led to the clinical issue. The safe decision therefore was to omit the outcome from both periods.

### *Dependent Variables*

Based on the research literature, the most relevant aspects of which were summarized above, the following five clinical outcomes were selected for this study: (1) conduction disorders; (2) cardiac dysrhythmia; (3) hypotension;[24] (4) new hip fractures; and (5) anticholinergic effects. These dependent variables are summarized below in Exhibit 5.9.2.

---

[21] This reaction, of course, may have been a decision to do nothing.

[22] The OPDUR system was in operation through June 1997. Complete data, however, was available only through 1996.

[23] Nursing home prescriptions that were batched and sent to demonstration pharmacies were not screened using OPDUR.

Conduction disorders and cardiac dysrhythmia are clear choices, although the ambiguity regarding mechanism of action is troubling. Orthostatic hypotension was also a natural choice, and hip fractures offer a clinically significant marker for that condition. Finally, the study analyzed a cluster of five anticholinergic effects (blurred vision, constipation, drowsiness, dry mouth, and urinary retention). This analysis is the most problematic because these effects may not be clinically significant enough to guarantee that physicians note them in patients' records or even that patients report them to physicians.

**Exhibit 5.9.2**

**HCFA DUR Demonstration Evaluation**
**Dependent Variables**

| Variable: | ICD-9 Classification: |
|---|---|
| conduction disorders | 426.0 - 426.9 |
| cardiac dysrhythmia | 427.0 - 427.9 |
| hypotension | 458.0 - 458.9 |
| new hip fractures | 820.0 - 820.9, also used CPT code |
| anticholinergic effects (any one of blurred vision, constipation, drowsiness, dry mouth, or urinary retention) | 368.8, 564.0, 780.09, 527.7, 788.2 |

*Sources:* International Classification of Diseases, 9th Revision, DHHS Publication No. (PHS) 91-1260, October, 1991.

Outcome indicators were constructed using Medicare diagnosis codes.[25] Each outcome was associated with a series of ICD-9 codes. If a patient's record showed evidence of one of those ICD-9s as either the primary or secondary diagnosis for an inpatient stay or for a physician encounter (paid by Part A or Part B of Medicare) during the follow-up period, then the patient was determined to have had the outcome.

For each outcome, a model also was estimated in which the outcome variable was based on inpatient claims alone. It is possible that compared to physicians in the outpatient setting, hospitals might be more precise in their coding of diagnoses because of its payment implications. Variables that were created using inpatient claims only therefore might contain less measurement error compared to variables created using all claims.

---

[24] Coding practice often does not distinguish between hypotension and orthostatic hypotension so a broad outcome encompassing either was created.

[25] Because all of the samples used in this analysis included individuals who were dually eligible for Medicare and Medicaid, it is assumed that all medical services were initially billed and primarily covered by Medicare.

In the new hip fracture model, the inpatient measures were particularly useful, because physician claims might capture follow-up care to a hip fracture rather than the hip fracture event itself. Hip fracture CPT codes for the main population were compared with inpatient claims for hip fracture to assure that the inpatient claims were capturing the hip fracture event. The rate of hip fracture based on CPT codes and the rate of hip fracture based on ICD-9 codes and inpatient claims were very similar. This similarity suggests that the hip fracture variable based on ICD-9 codes and inpatient claims is a good measure of the new hip fracture event.[26]

### *Independent Variables*

Each model included a basic set of covariates that captured exposure to the intervention as well as general predisposing and need characteristics. Additionally, covariates were added to each model to control for clinical risk factors that are specific to each of the outcomes. The main set of independent variables, as well as the outcome-specific independent variables, are described in the following sections.

***Exposure to the Intervention.*** Although this demonstration was designed as a random experiment, randomization occurred at the level of the pharmacy cluster while clinical benefits accrued at the level of the patient.[27] This posed several methodological challenges. First, individuals had the potential to go to multiple pharmacies and, potentially, to go to a mix of intervention, control, and non-demonstration pharmacies. The study adopted a simple approach. Individuals were considered members of the intervention group if they ever had a Medicaid-paid drug claim at an intervention pharmacy and never had such a claim at a control pharmacy. They were considered members of the control group if they ever had a Medicaid-paid drug claim at a control pharmacy and never had such a claim at an intervention pharmacy. Individuals who went to both intervention and control pharmacies were excluded from the study.[28]

***General Predisposing and Need Characteristics.*** The basic design of the study was to compare treatment and control patients within a regression framework that controls for demographic, drug, and disease factors. For all five outcomes, regressions included controls for the patient's race, sex, and age. These factors affect health outcomes both directly and as proxies for behavior; they may also affect the predisposition to use care. These variables, which were taken from the enrollment records of the Medicaid program, were included in regressions as a series of indicator variables.

---

[26] In the main population in Iowa, there were 396 hip fractures recorded using CPT code information, 81 hip fractures recorded using inpatient Medicare claims, and 110 hip fractures recorded using all Medicare claims. The higher rate of hip fractures recorded using all claims probably exists because outpatient, follow-up care for an old hip fracture has been coded as hip fracture.

[27] Pharmacy clusters, built around common prescribers, were created and randomized to minimize the influence that intervention pharmacies might exert through common prescribers on control patients. In Iowa, there were 191 pharmacy clusters ranging in size from 1 person to 141 people.

[28] For this purpose, the day that the patient first went to the demonstration pharmacy was included in the follow-up period.

*Length of Baseline and Follow-up Periods*

Other things equal, a patient with a longer follow-up period would be more likely to experience any clinical outcome than one with a shorter follow-up period, thus the length of the follow-up period was made an additional control in models. While the length of the baseline period does not affect patients' probabilities of clinical problems in and of itself, expected health spending, number of drugs, and risk factors will tend to be slightly under measured for patients with short baselines relative to patients with full baselines. Thus, controlling for measured covariates, the probability of clinical outcomes might be higher for patients with short baselines; consequently, the length of the baseline period was included as an additional control. The lengths of the follow-up and baseline period were entered in models as the natural log of the number of months in each period.

*Pharmacy Wave*

Pharmacies entered the cognitive services demonstration in two waves. The first wave of 119 pharmacies enrolled in June 1994. A second wave of 104 pharmacies entered in December 1994. It is possible that pharmacies entering the demonstration later learned from the experiences of prior entrants and used the OPDUR system more effectively. Indicator variables, therefore, were included in the models that capture the wave of the first demonstration pharmacy where the individual filled a prescription. The omitted category in the models was wave one.

The basic set of independent variables that were used in every model (as described above) are summarized below in Exhibit 5.9.3.

**Exhibit 5.9.3**

**HCFA DUR Demonstration Evaluation**
**Basic Set of Covariates Included in All Models**

race
sex
age (in categories)
length of baseline period
length of follow-up period
wave of first contact demonstration pharmacy

*Risk Factors Specific to Each Outcome*

*Conduction Disorders/Cardiac Dysrhythmia.* Analyses of these outcomes added additional controls for pre-existing conduction disorders, cardiac dysrhythmia, heart failure, and other cardiac diagnoses. These conditions are risk factors for ongoing conduction disorders and cardiac dysrhythmia. The models also controlled for the use of antiarrhythmics, beta blockers, CCBs, and digoxin in the baseline period because these drugs might also indicate pre-existing arrhythmia. Finally, analyses controlled for the use of TCAs during the baseline period; note that even if the main effect of the intervention were to take patients off TCAs, the change would be observed in the follow-up, not the baseline period.

***Hypotension.*** Analyses of this outcome controlled for pre-existing hypotension, the use of antihypertensives (which sometimes induce hypotension), TCAs, and low potency antipsychotics (also a risk factor) during the baseline period.

***Hip Fracture.*** Analyses of this outcome controlled for pre-existing hip fracture, hypotension, gait abnormality and osteoporosis. These conditions are known to raise the risk of hip fractures. In addition, the models controlled for the use of a number of drugs during the baseline: (1) sedative hypnotics or benzodiazepines, which may make patients drowsy and raise the risk of falls; (2) antihypertensives; (3) corticosteroids (chronic use) which potentially induce hypotension; and (4) TCAs.

***Anticholinergic Effects.*** Analyses of this outcome controlled for pre-existing anticholinergic effects, pre-existing diabetes (which may cause anticholinergic effects) and for incontinence and constipation (which show evidence of ongoing problems). The models also controlled for baseline use of a number of drugs known to induce anticholinergic effects: (1) antihistamines, (2) antispasmodics, (3) benzodiazepines, (4) narcotic analgesics, (5) low potency antipsychotics, and (6) TCAs.

### *Rate of Clinical Outcome in Pharmacy Cluster*

The pharmacy cluster, the unit of randomization in this analysis, effectively defined medical communities. Potentially, there could be unmeasured differences in demographic and behavioral characteristics among patients in different pharmacy clusters that influenced the need for care and the predisposition to seek it. Similarly, there could be geographic variation associated with cluster in access to care and in its practice, such as variations in coding of conditions and diagnoses and in treatment protocols. If these differences were associated with measured clinical outcomes, then controlling for them would improve the precision of estimated intervention effects. Thus, each model included the average outcome rate during the baseline period for each pharmacy cluster as an additional regressor. This variable served as a proxy for the net effect of geographic factors on the outcome of interest.

***Independent Variable Construction.*** The indicators for patients' baseline disease status were created in the same way as the indicators for patients' clinical outcomes. The appendix to this chapter shows the relevant ICD-9 codes. The indicators for patients' drug utilization were based on Medicaid drug claims during the baseline period. For all drugs the indicator was positive based on a single prescription claim.

## Models and Statistical Methods

The fact that pharmacy clusters, rather than individual patients, were randomized also meant that the standard regression framework required some adjustment. If indeed there were factors correlated with cluster that affected outcomes and if, despite the controls included in the models, these factors were not fully measured, then, because the number of clusters was fairly small (191) relative to the number of patients (4,220), these differences might not average out between intervention and control groups, despite the large number of patients, leading to misleading estimates of intervention effects. Specifically, traditionally-estimated standard errors would be smaller than the true standard errors.

As has been discussed, the first line of defense against this problem was the effort to include a full set of covariates, in particular the cluster averages discussed above. In addition, a cluster adjusted version of Huber/White's method was used to estimate the variance-covariance matrices of the models. This method relaxes the assumption of independence within clusters, and is considered to be a robust variance estimator in cases of cluster-based sampling.[29] An alternative approach would have been to bootstrap the coefficients of each model and derive nonparametric standard errors for the coefficients.[30] For the main population, this alternative method was tested; each antidepressant model was bootstrapped and standard errors were computed. The results, which are not presented here, were very similar to the results derived using the cluster-adjusted version of Huber/White variance estimator.[31]

Thus, the regression specification was:

$$y_{ijt_f}^{c} = \alpha + \beta_1 tr_{ij} + X_{it_b}\beta_2 + W_{it_b}^{c}\beta_3 + Z_i\beta_4 + \beta_5 y_{jt_b}^{c} + \nu_{ij}$$

with:

| | |
|---|---|
| $y^c_{ijtf}$ | indicator that patient i associated with pharmacy cluster j experienced clinical outcome c during the follow-up period (tf) |
| $tr_{ij}$ | indicator that patient i filled a prescription at treatment pharmacies and not at control pharmacies during the demonstration period |
| $X_{it}$ | general predisposing (socio-demographic) and need characteristics of patient i determined during the baseline period |
| $W^c_{itb}$ | patient i's risk factors for outcome c determined during the baseline period |
| $Z_i$ | length of follow-up and baseline windows for patient i |
| $y^c_{jtb}$ | rate of clinical outcome c in pharmacy cluster j during the baseline period |
| $v_{ij}$ | error |

---

[29] William Rogers originally developed this method in the early 1980's to account for cluster-based sampling issues in the RAND Health Insurance Experiment.

[30] The bootstrapping method involves pulling repeated samples (with replacement), estimating coefficients repeatedly (100 times in this case), and then deriving standard errors from the distributions of estimated coefficients. This method is often used in cases where we have no formulas to compute standard errors or when the assumptions we need to make to compute standard errors do not make sense (i.e. in this case, we could not assume that error terms were independently distributed because of inter-cluster correlation).

[31] Two other methods were considered to account for the effects of clustered sampling. The random-effects probit was considered, but the large size of the clusters made this model very difficult if not impossible to implement computationally. Generalized Estimating Equations (GEE) methods could have been used, but this approach would have made estimation more complex and the advantages of this approach were uncertain.

Because all dependent variables were binary, probit regressions were used. The probit coefficients were transformed to marginal effects to make interpretation easy.[32]

Exhibit 5.9.4 below summarizes all of the models that were estimated in this study. For all of the outcomes except new hip fractures, models were estimated using variables created from all Medicare claims and variables created from inpatient Medicare claims only. The new hip fracture models are based on inpatient claims only. Each model was estimated using the four populations described previously: (1) the main population of dually eligible beneficiaries; (2) a sub-population of dually eligible beneficiaries who used antidepressant drugs in the baseline or follow-up periods; (3) a sub-population of dually eligible beneficiaries who used antidepressant drugs in the baseline period only; and (4) a sub-sample of the main study population whose prescriptions triggered screen failures related to the use of antidrepressant drugs. Only models based on all claims and the main population are presented in full in Appendix G (these models are shaded in Exhibit 5.9.4).

## Findings

### *Descriptive Statistics*

Exhibit 5.9.5 below displays the means and standard deviations for the dependent variables and the basic set of covariates used in all of the models described above. The treatment group had a larger proportion of non-white individuals relative to the comparison group. Overall, however, the descriptive statistics reveal no striking evidence of a treatment effect — the treatment and comparison groups were quite similar along both baseline characteristics and follow-up outcomes.

### *Regression Results*

## Treatment Effects

Prospective drug use review, by alerting pharmacists to prescription problems at the point of service, might prevent adverse health outcomes found to be associated with antidepressant drug use. It was therefore hypothesized that the Iowa OPDUR system would be associated with lower rates of conduction disorders and cardiac dysrhythmia, two important adverse effects commonly associated with antidepressant use. The effect of the intervention was expected to be stronger among those with preexisting cardiac conditions because these individuals are at higher risk for cardiac problems related to antidepressant use.

---

[32] The marginal effect can be interpreted as the change in the probability of the outcome that is associated with an infinitesimal change in the independent variable, holding other covariates constant. For binary independent variables, the marginal effects are calculated based on a one unit change in the independent variable rather than an infinitesimal change.

Exhibit 5.9.4

**HCFA DUR Demonstration Evaluation Models Estimated**

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 1 | Conduction Disorders | Probit | All dual eligibles | 4220 | Based on all Medicare claims. |
| 2 | Conduction Disorders | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |
| 3 | Conduction Disorders | Probit | Antidepressant drug users in baseline or follow-up | 1042 | Based on all Medicare claims. |
| 4 | Conduction Disorders | Probit | Antidepressant drug users in baseline or follow-up | 1042 | Based on inpatient Medicare claims only. |
| 5 | Conduction Disorders | Probit | Antidepressant drug users in baseline | 690 | Based on all Medicare claims. |
| 6 | Conduction Disorders | Probit | Antidepressant drug users in baseline | 690 | Based on inpatient Medicare claims only. |
| 7 | Conduction Disorders | Probit | Those who triggered antidepressant drug screen failures | 331 | Based on all Medicare claims. |
| 8 | Conduction Disorders | Probit | Those who triggered antidepressant drug screen failures | 331 | Based on inpatient Medicare claims only. |
| 9 | Cardiac Dysrthymia | Probit | All dual eligibles | 4220 | Based on all Medicare claims. |
| 10 | Cardiac Dysrthymia | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |

Exhibit 5.9.4

## HCFA DUR Demonstration Evaluation Models Estimated

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 11 | Cardiac Dysrthymia | Probit | Antidepressant drug users in baseline or follow-up | 1042 | Based on all Medicare claims. |
| 12 | Cardiac Dysrthymia | Probit | Antidepressant drug users in baseline or follow-up | 1042 | Based on inpatient Medicare claims only. |
| 13 | Cardiac Dysrthymia | Probit | Antidepressant drug users in baseline | 690 | Based on all Medicare claims. |
| 14 | Cardiac Dysrthymia | Probit | Antidepressant drug users in baseline | 690 | Based on inpatient Medicare claims only. |
| 15 | Cardiac Dysrthymia | Probit | Those who triggered antidepressant drug screen failures | 331 | Based on all Medicare claims. |
| 16 | Cardiac Dysrthymia | Probit | Those who triggered antidepressant drug screen failures | 331 | Based on inpatient Medicare claims only. |
| 17 | Hypotension | Probit | All dual eligibles | 4220 | Based on all Medicare claims. |
| 18 | Hypotension | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |
| 19 | Hypotension | Probit | Antidepressant drug users in baseline or follow-up | 1042 | Based on all Medicare claims. |
| 20 | Hypotension | Probit | Antidepressant drug users in baseline or follow-up | 1042 | Based on inpatient Medicare claims only. |

**Exhibit 5.9.4**

## HCFA DUR Demonstration Evaluation Models Estimated

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 21 | Hypotension | Probit | Antidepressant drug users in baseline | 690 | Based on all Medicare claims. |
| 22 | Hypotension | Probit | Antidepressant drug users in baseline | 690 | Based on inpatient Medicare claims only. |
| 23 | Hypotension | Probit | Those who triggered antidepressant drug screen failures | 331 | Based on all Medicare claims. |
| 24 | Hypotension | Probit | Those who triggered antidepressant drug screen failures | 331 | Based on inpatient Medicare claims only. |
| 25 | New Hip Fracture | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |
| 26 | New Hip Fracture | Probit | Antidepressant drug users in baseline or follow-up | 1042 | Based on inpatient Medicare claims only. |
| 27 | New Hip Fracture | Probit | Antidepressant drug users in baseline | 690 | Based on inpatient Medicare claims only. |
| 28 | New Hip Fracture | Probit | Those who triggered antidepressant drug screen failures | 331 | Based on inpatient Medicare claims only. |
| 29 | Anticholinergic Effects | Probit | All dual eligibles | 4220 | Based on all Medicare claims. |
| 30 | Anticholinergic Effects | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |

**Exhibit 5.9.4**

## HCFA DUR Demonstration Evaluation Models Estimated

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 31 | Anticholinergic Effects | Probit | Antidepressant drug users in baseline or follow-up | 1042 | Based on all Medicare claims. |
| 32 | Anticholinergic Effects | Probit | Antidepressant drug users in baseline or follow-up | 1042 | Based on inpatient Medicare claims only. |
| 33 | Anticholinergic Effects | Probit | Antidepressant drug users in baseline | 690 | Based on all Medicare claims. |
| 34 | Anticholinergic Effects | Probit | Antidepressant drug users in baseline | 690 | Based on inpatient Medicare claims only. |
| 35 | Anticholinergic Effects | Probit | Those who triggered antidepressant drug screen failures | 331 | Based on all Medicare claims. |
| 36 | Anticholinergic Effects | Probit | Those who triggered antidepressant drug screen failures | 331 | Based on inpatient Medicare claims only. |

*Sources:* International Classification of Diseases, 9th Revision, DHHS Publication No. (PHS) 91-1260, October, 1991

**Exhibit 5.9.5**

## HCFA DUR Demonstration Evaluation
## Means and Standard Deviations of Dependent Variables and Covariates in Main Iowa Population Models
## (N = 4220)

| | Treatment (N = 1974) | Control (N = 2246) |
|---|---|---|
| Dependent Variables | | |
| conduction disorders | .056<br>(.229) | .057<br>(.233) |
| cardiac dysrhythmia | .240<br>(.427) | .240<br>(.427) |
| hypotension | .038<br>(.191) | .041<br>(.197) |
| new hip fracture | 020<br>(.141) | 018<br>(.134) |
| anticholinergic effects | .092<br>(.289) | .100<br>(.300) |
| Main Covariates | | |
| female | .820<br>(.384) | .823<br>(.382) |
| white*** | .866<br>(.340) | .934<br>(.249) |
| black*** | .057<br>(.232) | .034<br>(.182) |
| other race*** | .076<br>(.266) | .032<br>(.176) |
| age 65-69 | .213<br>(.410) | .204<br>(.403) |
| age 70-74 | .210<br>(.408) | .210<br>(.408) |
| age 75-79 | .160<br>(.366) | .178<br>(.383) |
| age 80-84 | .133<br>(.339) | .134<br>(.340) |
| age 85 or higher | .121<br>(.326) | .110<br>(.312) |
| length of baseline period | 10.9<br>(2.60) | 11.0<br>(2.49 |
| length of follow-up period | 17.9<br>(10.4) | 17.5<br>(10.1 |

Note: *** statistically significant difference at the .01 level
** statistically significant difference at the .05 level
* statistically significant difference at the .10 level

The estimation results, however, do not suggest the OPDUR system was effective in preventing conduction disorders or cardiac dysrhythmia. In the cardiac dysrhythmia models, the estimated treatment effect coefficient was generally inconsistent in sign and not statistically significant across all populations and models. Similar results were observed for estimates in the inpatient-only models. Treatment effect coefficients of the conduction disorders models also exhibited inconsistent signs between the main population and sub-population models. It is difficult, however, to interpret the fact that some of these coefficients were consistently positive and statistically significant among the sub-population models. One possible explanation is that there are other important covariates we have failed to incorporate in the regression models.

The OPDUR system also was expected to reduce the rate of hypotension and new hip fractures, the two most common adverse effects associated with antidepressant drug use. Holding differences between the treatment and control groups constant, however, the analyses did not reveal that there is a measurable treatment effect for these outcomes. Treatment effect coefficients were all statistically insignificant and with inconsistent signs across study populations and models.

Finally, if the OPDUR system is effective in screening inappropriate prescriptions, it should reduce the rate of anticholinergic outcomes among comparable patients. Indeed, the model results do suggest that, after controlling for the differences between the treatment and control groups, there is a statistically significant treatment effect in some of the sub-populations of antidepressant users. For patients who triggered antidepressant screen failures in the baseline period, OPDUR was associated with a eight percent reduction in the probability of anticholinergic effects in the follow-up period. In models based on inpatient claims only, baseline antidepressant users and those who triggered antidepressant screen failures, had a two percent and five percent reduction in the probabilities of anticholinergic effects respectively during the follow-up period. The treatment effect coefficients were all consistently negative although not statistically significant for the other sub-populations under study.

### Other Covariates

Among the demographic covariates, females were found to be less likely than males to experience cardiac dysrhythmia, hip fracture and anticholinergic effects, holding other factors constant. Compared to whites, blacks were less likely to have cardiac dysrhythmia, while individuals of other races were more likely to have cardiac dysrhythmia during the follow-up period. Otherwise, gender and race were not significant and consistent predictors of the other outcomes. Although there were no consistent relationship between all of the age categories and the outcomes, the two oldest sub-groups (80 to 85 and 85+ years) were on average more likely than younger patients to experience outcomes.

Baseline measures of diseases and drug utilization that were expected to place individuals at higher risk for outcomes were included as covariates in the regressions. In general, these covariates were excellent predictors of the outcomes of interest. The existence of cardiac conditions and the use of calcium channel blockers and digoxin during the baseline period, for example, were statistically significant predictors of conduction disorders and cardiac dysrhythmia. Baseline use of antimuscarinics, antispasmodics and antihistamines and baseline anticholinergic conditions were all associated with anticholinergic outcomes during the follow-up period.

**Exhibit 5.9.6**

**HCFA DUR Demonstration Evaluation**
**Summary of Signs of Treatment Effect Coefficients**
**Probit Models Based on Variables Created Using All Medicare Claims**

| **Population** | **Dependent Variable** | | | | |
|---|---|---|---|---|---|
| | **Conduction Disorders** | **Cardiac Dysrhythmia** | **Hypotension** | **New Hip Fracture** | **Anti-Cholinergic Effects** |
| Main | -.048<br>(.072) | -.005<br>(.012) | -.014<br>(.074) | .056<br>(.096) | -.072<br>(.057) |
| Baseline or Follow-up Users | .282**<br>(.141) | .071<br>(.093) | -.054<br>(.137) | .054<br>(.193) | -.095<br>(.109) |
| Baseline Users Only | .419**<br>(.198) | .096<br>(.118) | .015<br>(.179) | .141<br>(.259) | -.168<br>(.143) |
| Triggered Screen Failure | .639**<br>(.314) | .145<br>(.170) | -.201<br>(.256) | .294<br>(.404) | -.446**<br>(.206) |

Note:*** difference between arms statistically significant at the 1 percent level
** difference between arms statistically significant at the 5 percent level
* difference between arms statistically significant at the 10 percent level

*Sources:* Iowa Medicare claims

It was expected that the length of the baseline and follow-up periods might affect outcomes because individuals with longer baseline periods might have better information about baseline diseases and drug utilization, and because individuals with longer follow-up periods might have more time under observation. This hypothesis was confirmed for the follow-up length variable. Baseline length, however, was in general not a significant predictor of outcomes.

The wave variable were created to capture learning effects and other variation in the provision and efficiency of the interventions that might be correlated with the pharmacy's time of entrance into the demonstration. Pharmacies that entered the demonstration in the second wave, for example, might have learned from the experiences of pharmacies that entered during earlier wave. In the anticholinergic outcome model, we indeed found that pharmacies entered during the second wave had a lower rate of outcomes compared to the first wave pharmacies. For the other models, however, there was no statistically significant association between pharmacy wave and adverse outcomes related to antidepressant drug use.

## Pharmacy Cluster Effects

To account for the fact that clusters of pharmacies rather than individuals were randomized to the treatment and control groups, each model included as a covariate the average rate of the dependent variable in the individual's pharmacy cluster. It was expected that this variable would help control for unobservable, cluster level differences in factors that affect the outcomes of interest. The variable did turn out to be statistically significant and positive in sign in the conduction disorder outcome model, but, in general, the variables did not display a consistent pattern in sign across the models and they were often statistically insignificant. The Huber/White variance estimator, which also was used to adjust for cluster-based sampling, led to standard errors that were very similar to those calculated using a conventional variance estimator. Overall, then, adjustment for cluster-based sampling did not have any impact on the interpretation of the findings.

### Comment

The Iowa OPDUR system was designed to alert pharmacists to prescription problems at the point of service so that adverse outcomes could be avoided. This study focused on the impact of the OPDUR system on a set of clinical outcomes related to antidepressant drug use: (1) conduction disorders and dysrhythmia; (2) hypotension; and (3) anticholinergic effects. The analysis controlled for treatment and comparison group characteristics that may have differed at baseline, as well as the cluster-based sampling design used in the demonstration.

The findings indicate that the OPDUR system had a marginal impact in reducing the rate of anti-cholinergic outcomes. But, there was no measurable impact on the other outcomes of interest. There are a number of reasons why the impact might not exist, or why the effect might be too small to detect using the current research design. To begin with, for an OPDUR message to prevent an adverse outcome, a chain of events, potentially involving pharmacists, prescribers and patients, must take place. Assuming that the OPDUR system has correctly identified a potentially serious drug therapy problem, the pharmacist must then respond to the warning message in a timely manner that would reduce the risk of an adverse outcome. The pharmacist's actions may involve the prescriber and/or the patient. There are many possible ways in which an OPDUR message may not have led to this chain of events.

It is more likely, however, that limitations in the study design clouded measurement of an existing treatment effect, particularly if the treatment effect was very small. Generally, a cluster sampling approach reduces the power of detecting differences between the treatment and control groups. In this case, however, the intra-cluster correlation is likely to be low, and many of the clusters were very small; cluster-based sampling therefore may not seriously reduce the power of the study design.

More important is the fact that the treatment effect is indeed expected to be very small — it is hypothesized that the intervention might lead to half a percentage point difference in outcomes between the treatment and control groups. With the sample sizes used in the sub-population of baseline and follow-up period antidepressant drug users, the study probably had a less than ten percent power to detect a difference of this size. To capture such tiny differences between treatment and control groups, therefore, huge sample sizes would be needed.[33] Despite this limitation and other constraints of the study design, the findings offer no strong evidence that the Iowa OPDUR demonstration did lead to large declines in risk of adverse outcomes related to antidepressant drug use.

---

[33] The problem of inadequate power was anticipated early in the planning stages of this study. In an effort to increase sample sizes, dually eligible individuals from Maryland and Georgia who used antidepressants between 1993 and 1996 were considered as groups that could be used to augment the sample size. Since Maryland had an OPDUR system in operation between 1993 and 1996, Maryland antidepressant drug users could be added to the treatment group. Georgia, a state that never had an OPDUR system in place during this time period, was considered suitable as a proxy for the control group. The addition of Maryland and Georgia beneficiaries to this analysis would have had a sizeable

## 5.10 The Impact of the OPDUR Demonstration Project on Adverse Outcomes Related to the Use of Benzodiazepines

Prospective drug utilization review seeks to ensure rational drug therapy by identifying suboptimal therapies and intervening before medications are dispensed [Kralewski et al. (1990); Lipton and Bird (1991)]. By identifying significant therapeutic issues prior to dispensing, prospective drug utilization review clearly offers greater potential to prevent adverse consequences than does retrospective DUR.

As part of the Omnibus Budget Reconciliation Act of 1990 (OBRA '90), the Health Care Financing Administration funded a demonstration of on-line prospective drug utilization review (OPDUR) in Iowa. OPDUR uses computerized algorithms to screen prescriptions for therapeutic problems. Usually within seconds, and prior to dispensing, an OPDUR system alerts the pharmacist to drug therapy problems that have been identified in the prescription being screened.

The Iowa OPDUR system utilizes information about prescriptions filled at any pharmacy in the state that accepts Medicaid reimbursements; the statewide inter-pharmacy feature of the system allows the pharmacist to make more informed decisions that would be possible with a standard intra-pharmacy computer system. Furthermore, the Iowa OPDUR algorithms are based on the University of Maryland and Philadelphia College of Pharmacy and Sciences (UM/PCPS) criteria, which constitute the most refined set of criteria currently available. It was hypothesized therefore that the OPDUR system in Iowa would be effective in preventing drug therapy problems which have the potential to result in costly and potentially serious adverse health outcomes.

The Iowa OPDUR system was in operation between June 1993 and June 1997. Pharmacies entered the demonstration in two waves, and were randomly assigned to treatment and control groups in clusters. This chapter assessed the effects of the Iowa OPDUR demonstration on a set of adverse, clinical outcomes that have been linked to inappropriate prescribing of benzodiazepines. The population of interest in this study included elderly, Iowa residents who were eligible for both Medicare and Medicaid, and who filled a prescription at a demonstration pharmacy at least once.

---

impact on power; it was estimated that if dually eligible, antidepressant drug users from these states were added to the Iowa sample, the study would have 80 percent power to detect a treatment effect of size between .5 and 2 percent, depending on the outcome.

Incorporating Maryland and Georgia into the analysis has a number of potentially serious limitations, particularly regarding the fact that whether the two are really comparable study populations. If the Iowa analysis had led to the suggestive results, furthermore, the addition of Maryland and Georgia data to the analysis still would have been useful as a method of confirming the findings with larger sample sizes. The Iowa analysis in this chapter, however, did not offer any evidence that the OPDUR system was effective in reducing adverse outcomes related to antidepressant drug use. Consequently, it was decided that the disadvantages of adding data from other states to the study outweighed the advantages of greater statistical power.

If the OPDUR system was effective in improving the prescribing of benzodiazepines, then rates of falls and hip fractures would be expected to decline in the treatment group. As this chapter will show, holding other factors constant, treatment and control group patients generally did not have statistically significant differences in these outcome rates. The analysis presented in this chapter therefore indicates that the Iowa OPDUR system was not effective in preventing adverse outcomes associated with inappropriate prescribing of benzodiazepines.

## Adverse Clinical Outcomes Related to Use of Benzodiazepines

The elderly are well represented in the population of benzodiazepine users. Although the elderly account for only about 13 percent of the U.S. population they account for 35 to 40 percent of the prescriptions written for hypnotic agents [Gottlieb (1990)]. Clinical research has strongly suggested that the use and misuse of benzodiazepines is linked to adverse clinical effects among the elderly [Shorr and Robin (1996)]. The following sections provide a description of the outcomes selected for this study[34] and a brief review of the evidence for their association with benzodiazepine use.

### *Falls*

Falls and their consequences represent a major public health problem among the elderly. Studies on the incidence of accidental falls among the elderly suggest that on average, approximately 45 percent of community-dwelling persons over the age of 65 will fall at least once each year [Cumming and Klineberg (1994); O'Loughlin et al. (1993); Campbell, Borrie and Spears (1989)]. Falls are the leading accident occurring among the elderly, and are the leading cause of death from injury [Urton (1991); Sattin et al. (1990); Tinetti and Williams (1997)]. Among the elderly, the annual incidence of falls and fall-related hospital admissions increase sharply with advancing age [Steinweg (1997); Sattin et al. (1990); Lord (1990)].

The results of many studies support the conclusion that use of psychotropic drugs, especially benzodiazepines, is associated with risk for falling [Cumming (1998); Monane and Avorn (1996); Wickham, Cooper, Margetts and Barker (1989); Sorock and Shimkin (1988)]. In a recent prospective study of community-dwelling male veterans, Weiner, Hanlon and Studenski (1998) found that the risk of falls was significantly greater in CNS-active medication users as compared to non-users, and that fall liability increased with multiple CNS-active medication use. In an earlier study, Herings, Stricker, deBoer, Bakker, and Sturmans (1995) conducted a case-control study of persons age 55 and above, to examine the risk factor associated with falling leading to femur fractures. They found that falling was significantly associated with use of benzodiazepines, use of multiple benzodiazepines, and particularly with use of short half-life benzodiazepines. They concluded that the strong dose-response relationship with both short and long half-life benzodiazepines argues that the relationship between benzodiazepines with falls and femur fractures is probably more a function of high dosage than elimination half-life.

---

[34] In the planning phase for this study the rate of motor vehicle accidents had been identified as a clinical outcome that would be included in the benzodiazepine analyses. However, preliminary examination of the occurrence of moter vehicle accidents as encoded in the Medicare files indicated that there were too few cases (Main population, N = 30) to warrant further analyses.

*Hip Fractures*

With regard to hip fractures, the evidence for a link with psychotropic medications is mixed. A recent review of the epidemiology of hip fractures is provided by Cumming, Nevitt, and Cummings (1997). Two large population-based case control studies by Ray et al. (1987, 1989, 1991) examined the association of medications with hip fractures. The earliest of these studies [Ray et al. (1987)] linked prescription claims data with medical diagnoses of Medicaid beneficiaries. The results of this study indicated that antipsychotics, long-acting benzodiazepines, and tricyclics antidepressants were all significantly associated with increased risk for hip fracture. These finding were confirmed in a second study using claims data from the health insurance system in Saskatchewan, Canada [Ray et al. (1989, 1991)]. Opioid analgesics (codeine and/or propoxyphene) were also implicated in the second study as being associated with increased risk for hip fracture.

## Evaluation Approach

The objective of the present study was to examine the association between the Iowa OPDUR system and the occurrence of selected adverse events related to use of benzodiazepines. The expectation was that pharmacists using the OPDUR system would be more likely than pharmacists not using the OPDUR system to detect prescription problems related to benzodiazepine use and, consequently, to have an opportunity to prevent adverse outcomes. The occurrence of adverse clinical outcomes related to prescribing of benzodiazepines was therefore expected to decrease as a result of the safer prescribing and use of this class of prescription drugs.

The analytic strategy of this study was multivariate. For each candidate clinical outcome, the residual effect of the demonstration, after adjustment for known risk factors was examined. Multivariate regression procedures were used to control for demographic, disease, and drug factors predisposing a given patient to the outcome under study. All control variables were measured during a baseline period prior to the onset of the demonstration so as to be independent of any demonstration effects.

## Data

The data used in this study was extracted from Medicaid drug claims and Medicare health services utilization claims for patients in the state of Iowa in 1993 (to provide a baseline) and 1994-1996. (OPDUR was in operation in Iowa between June 1996 and June 1997.) Other records and data streams generated by the demonstration were also incorporated into the data set. Data development is described in detail in Chapter 4.7.

## Study Populations

Analyses examining the effect of the Iowa OPDUR system on benzodiazepine outcomes were conducted using four study populations, which are displayed in Exhibit 5.10.1. The four groups included: (1) a main study population; (2) a sub-sample of the main study population consisting of those who had been using benzodiazepines during the baseline period; (3) a sub-sample of the main population consisting of those who had been using benzodiazepines during the baseline or the follow-up period; and (4) a sub-sample of the study population whose prescriptions triggered screen failures

**Exhibit 5.10.1**

**HCFA DUR Demonstration Evaluation**
**Iowa OPDUR Evaluation**
**Study Populations**

| Study Population: | Description: |
|---|---|
| Main Population<br>(N = 4220) | Dually eligible for Medicare and Medicaid, at least 65 years old, not enrolled in a HMO or living in a nursing home, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies |
| Baseline and Follow-up Benzodiazepine Users<br>(N = 1174) | Dually eligible for Medicare and Medicaid, at least 65 years old, not enrolled in an HMO or living in a nursing home, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled at least one prescription for a benzodiazepine during the baseline period or the follow-up period |
| Baseline Benzodiazepine Users<br>(N = 799) | Dually eligible for Medicare and Medicaid, at least 65 years old, not enrolled in a HMO or living in a nursing, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled at least one prescription for a benzodiazepine during the baseline period |
| Triggered a Benzodiazepine Screen Failure<br>(N = 764) | Dually eligible for Medicare and Medicaid, at least 65 years old, not enrolled in a HMO or living in a nursing home, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, triggered at least one screen failure related to benzodiazepine use |

*Sources:* IA Medicaid claims

related to the use of benzodiazepines. Although all models were tested using these four samples, not all results are presented in this chapter. Rather, this chapter focuses on interpreting and summarizing the analytic results. Complete model results for selected models using the main population are presented in full in the appendix to this chapter.

The main study population consisted of 4220 elderly patients (65 and over on January 1, 1993) who were eligible for both the Medicaid and Medicare programs, who were not enrolled in a HMO, not living in a nursing home, and who went to either intervention or control pharmacies during the intervention. The requirement of dual eligibility ensures that the analyses have complete and consistent records of drug and health services utilization for all patients in the data set.

The study population was further limited to individuals who used benzodiazepines during the baseline period, and individuals who used benzodiazepines during the baseline or follow-up periods. Narrowing the population to the elderly was intended to improve the specificity of the analysis. Furthermore, the analysis considered both a sub-sample of baseline benzodiazepine users as well as a sub-sample of baseline or follow-up benzodiazepine users because of concerns that the intervention might affect benzodiazepine use in the follow-up period. In this case, estimates based on the baseline/follow-up sub-sample might be biased.

The models also were estimated using a sub-sample that consisted of those patients whose prescriptions had triggered at least one screen failure. In the treatment group, the pharmacist received the screen failure information on his/her computer and presumably read and reacted to the information. In the control group, screen failure information was recorded for research purposes only and was not transmitted to the pharmacist. In the sub-population of persons who triggered screen failures, therefore, this analysis compares the rate of adverse outcomes between individuals whose pharmacists were alerted via OPDUR to a benzodiazepine prescription problem and individuals whose pharmacists were not alerted via OPDUR to a benzodiazepine prescription problem. The screen trigger sample is the most targeted in the sense that every individual in the sample had a benzodiazepine prescribing problem; compared to the other samples analyzed, this group was at highest risk of adverse effects. However, the sample size is reduced considerably by excluding those who did not trigger a screen failure.

### Definitions of Baseline and Follow-up Periods

Pharmacies enrolled in the Iowa OPDUR demonstration in two waves (June 1994 and December 1994). This study assumes that a patient's exposure to the demonstration (and consequently the patient's follow-up period) began when a patient first filled a prescription at a pharmacy participating in the demonstration. In the majority of cases, the follow-up period then continued until the last month of 1996. If a patient died prior to the end of the study, their follow-up period ended on the date of death. In other cases, if the patient lost eligibility for Medicare or Medicaid, entered a nursing home, or entered an HMO, the follow-up period also ended sooner because the outcomes data represented in the claims would be incomplete.

In addition to variation in follow-up periods, patients' baseline periods could also differ in length. The baseline period was determined to end on the day before the patient's first contact with a demonstration pharmacy and to begin either twelve months prior to the ending date or on the date that the patient first became eligible for the Medicare program, became eligible for the Medicaid program, left an HMO, or left a nursing home and became resident in the community. Similar to the restrictions on follow-up period, the baseline period restrictions were imposed in order to ensure that the medical record was complete during the baseline period.

Clinical outcomes that occurred on the day the patient first went to a demonstration pharmacy are excluded from *both* the baseline and the follow-up periods, because in such cases, it is impossible to know whether the clinical issue precipitated the prescription or the prescription led to the clinical issue. The conservative decision was to omit the outcome from both the baseline and follow-up periods.

### Dependent Variables

The outcomes selected for this study were falls and hip fractures. Outcome indicators were constructed using Medicare diagnosis codes.[35] Each outcome was associated with a series of ICD-9 codes. For falls, if a patient's record showed evidence of one of those ICD-9s as either the primary or secondary diagnosis for an inpatient stay or for a physician encounter (paid by Part A or Part B of Medicare) during the follow-up period, then the patient was determined to have experienced the event. A model also was estimated in which the falls variable was based on inpatient claims alone. The rationale for this is based on the possibility of differential validity of the outcome diagnoses. It is possible that, relative to physicians in the outpatient setting, hospitals may be more precise in their coding of diagnoses because of payment considerations. Variables that were created using inpatient claims only may thus contain less measurement error compared to variables created using all claims.

For hip fractures, index variables were computed using records of inpatient stays and physician encounters and from inpatient stays only. Further, an indicator variable based on verification of hip fractures via CPT codes was computed. Preliminary analyses for the main population showed good correspondence between the index variable based on inpatient records and the CPT verified events. It was felt that the higher rate of hip fractures observed using all Medicare claims probably exists because outpatient, follow-up care, for old hip fractures has been coded as new hip fractures. For this study, new hip fractures based on inpatient claims was selected as the appropriate outcomes event. Exhibit 5.10.2 presents the associated ICD-9 codes used to define the indicator variables for these outcomes.

**Exhibit 5.10.2**

**HCFA DUR Demonstration Evaluation**
**Benzodiazepine Outcomes**
**OPDUR Evaluation, Dependent Variables**

| Variable: | ICD-9 Classification: |
|---|---|
| Fall Occurrence | E880-E888 |
| New Hip Fractures | 820.0-820.9, confirmed by CPT Codes |

*Sources:* International Classification of Diseases, 9th Revision, DHHS Publication No. (PHS) 91-1260, October, 1991.

Outcome indicators were constructed using Medicare diagnosis codes. Each outcome was associated with a series of ICD-9 codes. If a patient's record showed evidence of one of those ICD-9s as either the primary or secondary diagnosis for an inpatient stay or for a physician encounter (paid by Part A or

---

[35] Because all of the samples used in this analysis included individuals who were dually eligible for Medicare and Medicaid, it is assumed that all medical services were initially billed and primarily covered by Medicare.

Part B of Medicare) during the follow-up period, then the patient was determined to have had the outcome.

For falls, a model also was estimated in which the outcome variable was based on inpatient claims alone. It is possible that compared to physicians in the outpatient setting, hospitals might be more precise in their coding of diagnoses because of its payment implications. Variables that were created using inpatient claims only therefore might contain less measurement error compared to variables created using all claims. For falls, models based on variables created using inpatient claims only are not presented here because the results were very similar to models based on variables created using all claims.

## Independent Variables

### *Basic Covariates*

Each model included a basic set of covariates that captured exposure to the intervention as well as general predisposing and need characteristics. Additionally, covariates were added to each model to control for clinical risk factors specific to each of the outcomes. The main set of independent variables, as well as the outcome-specific independent variables, are described in the following sections.

## Exposure to the Intervention

Although the OPDUR demonstration was designed as a random experiment, randomization occurred at the level of the pharmacy cluster while clinical benefits were expected to accrue at the level of the patient.[36] Individuals had the potential to fill prescriptions at different pharmacies and, potentially, to fill prescriptions at a mix of intervention, control, and non-demonstration pharmacies. It could be expected that the effect of the demonstration would be diluted for such individuals depending on the relative proportions of prescriptions filled at the different pharmacy types. To resolve this methodological problem, this study adopted a simple approach. Individuals were considered members of the intervention group if they ever had a Medicaid-paid drug claim at an intervention pharmacy and never had such a claim at a control pharmacy. Similarly, they were considered members of the control group if they ever had a Medicaid-paid drug claim at a control pharmacy and never had such a claim at an intervention pharmacy. Individuals who went to both intervention and control pharmacies were excluded from the study regardless of the number of filled prescription at either intervention or control pharmacies.[37]

General Predisposing and Need Characteristics – The basic analytic strategy for the several clinical outcomes studies reported here was to examine the effect of the treatment variable within a regression framework that statistically controlled for demographic, drug use, and disease factors. For all outcomes, the regression models included variables for the patient's race, sex, and age. These factors affect health outcomes indirectly as proxies for health status. They may also affect the probability of having access to, or using, health care services. Thus these measures serve to control for general predisposing and need characteristics. In this study on benzodiazepine-related clinical outcomes, age, race, and sex are also known to act directly as risk factors to the outcomes of interest.

---

[36] Pharmacy clusters, built around common prescribers, were created and randomized to minimize the influence that intervention pharmacies might exert through common prescribers on control patients.

[37] For this purpose, the day that the patient first went to the demonstration pharmacy was included in the follow-up period.

### Length of Baseline and Follow-up Periods

It would be expected that patients with a longer follow-up period would be more likely to experience any given clinical outcome than one with a shorter follow-up period. Therefore, the length of the follow-up period was included as a covariate in the models. While the length of the baseline period does not affect patients' probabilities of experiencing a specific clinical problem, it is likely that the assessment of health expenditures, number of drugs, and risk factors will tend to be measured with less precision for patients with short baseline periods relative to patients with full baseline periods. Thus, controlling for other measured covariates, the probability of clinical outcomes might be higher for patients with short baselines; consequently, the length of the baseline period was included as an additional covariate. The lengths of the follow-up and baseline period were entered into all models as the natural log of the number of months in each period.

### Pharmacy Wave

Pharmacies entered the OPDUR demonstration in two waves. The first wave of 119 pharmacies enrolled in June 1994. A second wave of 104 pharmacies of pharmacies entered in December 1994. It is possible that pharmacies entering the demonstration later learned from the experiences of prior entrants and used the OPDUR system more effectively. For these reasons, indicator variables were included in the models to assess the impact of the wave of the first demonstration pharmacy where the individual filled a prescription. For all models, the reference category was wave one. The basic set of independent variables that were used in every model (as described above) are summarized below in Exhibit 5.10.3.

**Exhibit 5.10.3**

**HCFA DUR Demonstration Evaluation**
**Basic Set of Covariates Included in All Models**

race
sex
age (in categories)
length of baseline period
length of follow-up period
wave of first contact with demonstration pharmacy

### Covariate Risk Factors Specific to Each Outcome

In addition to the basic set of independent variables each regression model included measures of risk factors specific to each outcome.

### Falls

Risk factors for falls are many and varied, including physiological impairments, acute or chronic diseases, and medications that impair sensory, cognitive, or musculoskeletal functioning [Cumming (1998)]. The evidence for multiple concurrent prescription drug use as an independent risk factor for falling is fairly strong. In particular, many studies have identified psychotropic medications to be risk factors for falls. There is also evidence implicating selected cardiovascular, analgesic, and other classes of medications as risk factors for falls through their direct and indirect effects on psychomotor functioning and balance [Cumming (1998); Monane and Avorn (1996)]. Since the number of specific drugs that have been implicated for falls is large, and also because the primary focus of the study is on the effect of the intervention rather than drug use per se, it was deemed appropriate to reduce the number of covariates by developing index variables based on clusters of medication use. For the falls outcome, three clusters were defined, psychotropics and related, cardiovascular and related, and other risk-related medications. Exhibit 5.10.4 presents the set of risk factors included in analyses involving falls.

### Hip Fractures

Diseases, conditions, and medications that affect bone density have been repeatedly identified in numerous studies as risk factors for hip fractures. Additionally, those diseases, conditions, and medications that predispose to falls are also associated with risk for hip fractures. As indicated above, most studies have identified multiple medications to be positively associated with falls. For hip fractures, however, there is some evidence that thiazide diuretics may be protective due to their reduction of calcium excretion. For this reason, use of thiazide diuretics was retained as an single independent covariate whereas the other medications known to be associated with hip fracture risk were aggregated as described above for the falls outcome. Exhibit 5.10.5 present the covariates included in models that examine the effect of the intervention on the hip fracture outcome..

### Rate of Clinical Outcome in Pharmacy Cluster

The pharmacy cluster, which is the unit of randomization in this analysis, effectively defined medical communities. Potentially, there could be unmeasured differences in demographic and behavioral characteristics among patients in different pharmacy clusters that influenced the need for care and the predisposition to seek it. Similarly, there could be geographic variation associated with individuals cluster with respect to access to care and in medical practice. For example, clusters might vary in accepted treatment protocols, or in coding of specific conditions and diagnoses. If these differences were associated with measured clinical outcomes, then controlling for them would improve the precision of estimated intervention effects. In order to control for such differences, each model included the average outcome rate during the baseline period for each pharmacy cluster as an additional regressor. This variable served as a proxy measure for the net effect of geographic factors on the outcome of interest.

**Exhibit 5.10.4**

### HCFA DUR Demonstration Evaluation
### Disease, Condition, and Drug Use Covariates for Models Involving Falls as the Dependent Variable

Disease/Condition Covariates

- Arthritis
- Convulsions
- Dementia
- Depression
- Diabetes
- Epilepsy
- Falls
- Gait abnormality
- Hip fracture
- Hypotension
- Orthostatic hypotension
- Syncope and collapse

Medication Use Covariates

- Psychotropics and related medications
    - Antidepressants
    - Antipsychotics
    - Anticonvulsants
    - Non-benzodiazepine CNS depressants

- Cardiovascular and related medications
    - Antiarrythmics
    - Beta blockers
    - Calcium channel blockers
    - Digoxin

- Other medications related to risk of falls
    - Aminoglycosides
    - Antihistamines
    - Narcotic analgesics

Multiple medications

**Exhibit 5.10.5**

## HCFA DUR Demonstration Evaluation
## Disease, Condition, and Drug Use Covariates for Models Involving Falls as the Dependent Variable

Disease/Condition Covariates
- Arthritis
- Dementia
- Diabetes
- Epilepsy
- Falls
- Gait Abnormality
- Hip Fracture
- Hypotension
- Orthostatic hypotension
- Osteoporosis
- Osteomalacia
- Syncope and Collapse

Medication Use/Covariates
- Psychotropics and related
  - Antidepressants
  - Antipsychotics
  - Anticonvulsants
  - Non-benzodiazepine CNS depressants

- Cardiovascular and related
  - Antiarrythmics
  - Beta Blockers
  - Calcium Channel Blockers
  - Digoxin

- Other medications related to increased risk of hip fracture
  - Aminoglycosides
  - Antihistamines
  - Corticosteroids
  - Estrogens
  - Narcotic analgesics

- Thiazide diuretics

- Multiple prescriptions

### Independent Variable Construction

The indicators for each patient's baseline disease and drug use status were created in the same way as the indicators for patients' clinical outcomes. Specifically, indicators for diseases and conditions were based on diagnosis codes for Medicare claims during the baseline period. As previously noted, the prior hip fracture indicator was based on inpatient claims only. The indicators for patients' drug utilization were based on Medicaid drug claims during the baseline period. For all drugs, the usage indicator was positive if an individual had one or more prescription claims for that drug or drug class.

## Models and Statistical Methods

Because all dependent variables were categorical and binary, probit regressions were used. The regression specification for the examination of treatment effects in this study was:

$$y_{ijt_f}^{c}=\alpha+\beta_1 tr_{ij}+X_{it_b}\beta_2+W_{it_b}^{c}\beta_3+Z_i\beta_4+\beta_5 y_{jt_b}^{c}+\nu_{ij}$$

with:

| | |
|---|---|
| $y^{c}_{ijtf}$ | indicator that patient i associated with pharmacy cluster j experienced clinical outcome c during the follow-up period (tf) |
| $tr_{ij}$ | indicator that patient i filled a prescription at treatment pharmacies and not at control pharmacies during the demonstration period |
| $X_{it}$ | general predisposing (socio-demographic) and need characteristics of patient i determined during the baseline period |
| $W^{c}_{itb}$ | patient i's risk factors for outcome c determined during the baseline period |
| $Z_i$ | length of follow-up and baseline windows for patient i |
| $y^{c}_{jtb}$ | rate of clinical outcome c in pharmacy cluster j during the baseline period |
| $\nu_{ij}$ | error |

Exhibit 5.10.6 below summarizes all of the models that were estimated in this study. For falls, models were estimated using variables created from all Medicare claims and variables created from inpatient Medicare claims only. The new hip fracture models are based on inpatient claims only. Each model was estimated using the four populations described previously: (1) the main population of dually eligible beneficiaries; (2) a sub-population of dually eligible beneficiaries who used benzodiazepines in the baseline or follow-up periods; (3) a sub-population of dually eligible beneficiaries who used benzodiazepines in the baseline period only; and (4) a sub-population of dually eligible beneficiaries whose prescriptions triggered screen failures related to the use of benzodiazepines. Only models based on all claims and the main population are presented in full in Appendix H (these models are shaded in Exhibit 5.10.6).

**Exhibit 5.10.6**

**HCFA DUR Demonstration Evaluation**
**Models Estimated**

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 1 | Falls | Probit | All dual eligibles | 4220 | Based on all Medicare claims. |
| 2 | Falls | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |
| 3 | Falls | Probit | Benzodiazepine users in baseline | 799 | Based on all Medicare claims. |
| 4 | Falls | Probit | Benzodiazepine users in baseline | 799 | Based on inpatient Medicare claims only. |
| 5 | Falls | Probit | Benzodiazepine users in baseline or follow-up | 1174 | Based on all Medicare claims. |
| 6 | Falls | Probit | Benzodiazepine users in baseline or follow-up | 1174 | Based on inpatient Medicare claims only. |
| 7 | Falls | Probit | Benzodiazepine users who triggered screen failures | 764 | Based on all Medicare claims. |
| 8 | Falls | Probit | Benzodiazepine users who triggered screen failures | 764 | Based on inpatient Medicare claims only. |
| 9 | Hip Fractures | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |
| 10 | Hip Fractures | Probit | Benzodiazepine users in baseline | 799 | Based on inpatient Medicare claims only. |

**Exhibit 5.10.6**

**HCFA DUR Demonstration Evaluation**
**Models Estimated**

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 11 | Hip Fractures | Probit | Benzodiazepine users in baseline or follow-up | 1174 | Based on inpatient Medicare claims only. |
| 12 | Hip Fractures | Probit | Benzodiazepine users who triggered screen failures | 764 | Based on inpatient Medicare claims only. |

## Findings

### *Descriptive Statistics*

Exhibit 5.10.7 below displays the means and standard deviations for the dependent variables and the basic set of covariates used in all of the models described above. The treatment group had a larger proportion of non-white individuals relative to the comparison group. Additionally, the treatment group had higher representation of wave 1 pharmacies compared to the control group, which had more wave 2 representation. Aside from these differences, the treatment and comparison groups were quite similar with respect to the main covariates. Comparison of the group means for the dependent variables involving falls and hip fractures reveals no striking evidence of a treatment effect for the outcome measures. For all three dependent variables, the treatment and comparison groups were quite similar.

### *Treatment Effects*

It was hypothesized that individuals who visited pharmacies participating in the OPDUR system would be less likely to be exposed to inappropriate prescribing of benzodiazepines, which in turn would lead to reduced risk of adverse outcomes related to benzodiazepine use. This analysis focused on two preventable adverse outcomes related to benzodiazepines use: (1) falls; and (2) hip fractures. If the feedback to pharmacists provided by the OPDUR system did in fact significantly decrease rates of these problems, the models should indicate that, after adjusting for baseline demographic and risk factors and for factors related to potential design effects, the treatment group should have had lower rates of adverse benzodiazepine effects relative to the control group. If this were the case, the probit coefficients should be negative and significant.

**Exhibit 5.10.7**

### HCFA DUR Demonstration Evaluation
### Means and Standard Deviations of Dependent Variables and Covariates in Main Iowa Population Models
### (N=4220)

| | Treatment (N=1977) | Control (N=2243) |
|---|---|---|
| Dependent Variables | | |
| falls (inpatient/outpatient) | 0.0526<br>(0.2233) | 0.0602<br>(0.2379) |
| falls (inpatient) | 0.0309<br>(0.1730) | 0.0321<br>(0.1763) |
| new hip fracture (inpatient) | 0.0202<br>(0.1408) | 0.0183<br>(0.1340) |
| Main Covariates | | |
| female | 0.8199<br>(0.3843) | 0.8230<br>(0.3817) |
| white*** | 0.8665<br>(0.3402) | 0.9336<br>(0.2491) |
| black*** | 0.0572<br>(0.2322) | 0.0343<br>(0.1821) |
| other race*** | 0.0764<br>(0.2657) | 0.0321<br>(0.1763) |
| age 65-69 | 0.2696<br>(0.4439) | 0.2523<br>(0.4345) |
| age 70-74 | 0.2559<br>(0.4365) | 0.2572<br>(0.4372) |
| age 75-79 | 0.2003<br>(0.4003) | 0.2167<br>(0.4121) |
| age 80-84 | 0.1527<br>(0.3598) | 0.1645<br>(0.3708) |
| age 85 or higher | 0.1214<br>(0.3267) | 0.1092<br>(0.3120) |
| length of baseline period | 10.9029<br>(2.5996) | 11.0134<br>(2.4919) |
| length of follow-up period | 17.8852<br>(10.3588) | 17.5412<br>(10.0900) |
| wave 1 pharmacy*** | 0.6419<br>(0.4796) | 0.4717<br>(0.4993) |
| wave 2 pharmacy*** | 0.3581<br>(0.4796) | 0.5283<br>(0.4993) |

Note
*** statistically significant difference at the .01 level
** statistically significant difference at the .05 level
* statistically significant difference at the .10 level

Exhibits 5.10.9 in the Appendix provide the coefficients and their respective standard errors for all variables included in the analysis of falls based on the main population. Analogous results for new hip fractures are provided in Appendix J. The probit results for the treatment factor for all 12 models[38] are summarized in Exhibit 5.10.8.

**Exhibit 5.10.8**

**HCFA DUR Demonstration Evaluation**
**Summary of Treatment Effects: Probit Coefficients**

| | Dependent Variable | | |
|---|---|---|---|
| **Population** | **Falls, all Medicare records** | **Falls, inpatient Medicare records only** | **Hip fractures, inpatient Medicare records only** |
| Main | -.103<br>(.067) | -.037<br>(.081) | .455<br>(.097) |
| Baseline Benzodiazepine Users Only | -.135<br>(.116) | -.056<br>(.174) | -.102<br>(.212) |
| Baseline and Followup Benzodiazepine Users | -.095<br>(.142) | -.042<br>(.143) | -.129<br>(.175) |
| Triggered a Benzodiazepine Screen Failure | -.032<br>(.163) | -.240<br>(.200) | -.309<br>(.217) |

Note: *** difference between arms statistically significant at the 1 percent level
** difference between arms statistically significant at the 5 percent level
* difference between arms statistically significant at the 10 percent level
Source: IA Medicare claims

Examination of the estimation results summarized in Exhibit 5.10.8 suggests that the Iowa OPDUR system was not effective in preventing either of fall or hip fracture rates. Although 11 of the 12 treatment coefficients are negative in sign, which would be consistent with a protective treatment effect, none of these results are statistically significant at the .05 level. These findings suggest that the OPDUR system was ineffective in preventing both falls and hip fractures related to inappropriate prescribing of benzodiazepines.

### *Other Covariates*

Gender and race generally were not statistically significant predictors of falls and hip fractures, although the main sample model for falls using all Medicare data indicated that race other than White or Black was associated with increased risk of falling (p=0.02). For most models, increasing age was associated with increased incidence of adverse outcomes, with individuals in the oldest age groups experiencing the highest incidence of falls and hip fractures.

Baseline measures of diseases and drug utilization that were expected to place individuals at higher risk for outcomes were included as covariates in the models. In general, covariates measuring baseline disease conditions were poor predictors of the outcomes. Exceptions included convulsions, which was positively associated with falls using all Medicare data, and arthritis, which was negatively associated with hip fractures in all four samples. The baseline measures of drug utilization for drug classes expected to affect risk of benzodiazepine-related adverse outcomes also proved to be inconsistent predictors. For example, models for outcomes involving falls based on all Medicare claims data did not reveal any significant associations of the drug class indicators with the outcome. For outcomes involving falls based only on inpatient data, there were only two associations which were significant at 0.10, both of which occurred in the screen trigger group. For this sample, the use of cardiac drugs hypothesized to increase fall risk was marginally associated with decreased risk of falls (p=0.06), while use of other medications was associated with increased risk (p=0.08). Similarly, for the outcome of hip fractures, no drug class variable was significant at the 0.05 level, and only two associations had significance levels below 0.10. In the sample of baseline or followup benzodiazepine users, individuals who used thiazide diuretics had a lower incidence of hip fracture (p=0.06). In the sample of benzodiazepine users who triggered screen failures, use of psychotropic drugs hypothesized to increase risk of hip fracture was marginally associated with higher risk (p=0.09). The only other drug variable which emerged in the benzodiazepine models was the log of the maximum generic products utilized in a single baseline month, which was a measure of baseline polypharmacy. This variable was strongly and positively associated with falls based on all Medicare data for the main sample (p=.001) and to a lesser degree in the sample of baseline or followup users (p=0.10).

---

[38] For some sub-populations it was not possible to include all disease covariates as risk factors due to lack of adequate sample size. For these situations, the risk factor was omitted and the model was re-fitted.

It was expected that the length of the baseline and follow-up periods might affect outcomes because individuals with longer baseline periods might have better information about baseline diseases and drug utilization, and because individuals with longer follow-up periods might have more time under observation. This hypothesis was confirmed for the follow-up length variable seven of the eight fall models; baseline length, however, was not a significant predictor of falls. Neither baseline nor follow-up period emerged as a significant predictor of hip fractures.

The wave variable was created to capture learning effects and other variation in the provision and efficiency of OPDUR interventions that might be correlated with the pharmacy's date of entrance into the demonstration. In the sample of baseline benzodiazepine users, the model results suggest that individuals visiting pharmacies that enrolled during wave two were less likely to experience hip fractures relative to those visiting pharmacies that enrolled during the first wave (p=0.06). No other significant associations involving the timing of pharmacy entrance into the demonstration were found.

### *Pharmacy Cluster Effects*

To account for the fact that clusters of pharmacies rather than individuals were randomized to the treatment and control groups, each model included as a covariate the average rate of the dependent variable in the individual's pharmacy cluster. It was expected that this variable would help control for unobservable, cluster level differences in factors that affect the outcomes of interest. These variables were not consistently statistically significant across the models. Only one model revealed a marginally significant effect for pharmacy cluster: for hip fractures in the sample of baseline benzodiazepine users, the baseline rate of occurrence of hip fractures within the pharmacy cluster was associated with increased risk of hip fracture (p=0.08). No other significant effects of pharmacy cluster were found. Overall, then, it can be concluded that adjustment for cluster-based sampling did not have any impact on the interpretation of the findings for adverse outcomes associated with benzodiazepines.

## Comment

The Iowa OPDUR system was designed to alert pharmacists to prescription problems at the point of service so that the risk for adverse outcomes could be reduced or avoided. This chapter focused on the impact of the intervention on two clinical outcomes related to benzodiazepine use: falls and hip fractures. The analysis controlled for treatment and comparison group characteristics that may have differed at baseline, as well as the cluster-based sampling design used in the demonstration. Based on the regression results discussed above, there is no evidence that the OPDUR alerts related to benzodiazepine therapy prevented either of the two adverse outcomes that expected to be responsive to the intervention.

There are several reasons that may offered to explain why no significant effect was found. Because previously-validated criteria were used in the OPDUR system, it is unlikely that the OPDUR system failed to identify potentially serious drug therapy problems. However, it may be that the risk situations that were detected were not able to be greatly ameliorated due to the importance of continuing the clinical care for which the drugs were prescribed.

Furthermore, even if changes in drug therapy were possible, in order for a reduction in outcomes to occur, a specific chain of risk-reducing actions must take place once a screen failure has been generated. These actions usually require a coordinated set of responses on the part of the pharmacist, the prescriber, and the patient. The failure of the OPDUR system to produce observable reductions in falls and fractures may be a reflection of a systematic failure of one or more of the responsible parties to accept and act on the OPDUR alert.

It is also possible that limitations in the study design complicated the measurement and detection of existing treatment effects, particularly if the treatment effects were very small. Generally, a cluster sampling approach reduces the power of detecting differences between the treatment and control groups. In this case, however, the intracluster correlation is likely to be low, and many of the clusters were very small; cluster based sampling therefore does not seriously reduce the power of the study design.

Another important fator is that the treatment effect is expected to be very small. With the sample sizes used and the low observed incidence of falls and fractures in the sub-population of baseline and follow-up period benzodiazepine drug users, the study had somewhat limited power to detect small differences between treatment and control groups. To assess small differences between treatment and control groups huge sample sizes would be needed. Although the sample size constraints may have limited the ability of the study to detect significant effects, the findings nevertheless suggest that the Iowa OPDUR demonstration did not lead to large enough effects to be observed without multivariate analyses. Moreover, attempts to detect even a small effect via statistical control of many risk factors were unsuccessful.

## 5.11 The Impact of the Iowa OPDUR Demonstration Project on Adverse Outcomes Related to Inappropriate Prescribing of Cardiovascular Drugs

Inappropriate prescribing can lead to costly, serious adverse health outcomes. Prospective drug use review is intended to prevent these adverse health outcomes by alerting pharamacists to drug therapy problems at the time the prescription is filled. Unlike retrospective drug use review (RDUR), which is conducted after the prescription has been dispensed, prospective drug use review has the potential to resolve drug therapy problems before they result in illness or even death [Chrischilles et. al. (1997)].

As part of the Omnibus Budget Reconciliation Act of 1990 (OBRA '90), the Health Care Financing Administration funded a demonstration of on-line prospective drug utilization review (OPDUR) in Iowa. OPDUR is a form of prospective drug use review that relies on computerized algorithms to screen prescriptions for problems. Usually within seconds, and at the point of service, an OPDUR system alerts the pharmacist to drug therapy problems that have been identified in the prescription being screened.

The Iowa OPDUR system utilizes information about prescriptions filled at any pharmacy in the state that accepts Medicaid; this feature of the system allows the pharmacist to make more informed decisions that would be possible with a standard pharmacy computer system. Furthermore, the Iowa OPDUR algorithms are based on the University of Maryland and Philadelphia College of Pharmacy and Sciences (UM/PCPS) criteria, which constitute the most refined set of criteria currently available. It was hypothesized therefore that the OPDUR system in Iowa would be effective in preventing drug therapy problems which have the potential to result in costly and potentially serious adverse health outcomes.

The Iowa OPDUR system was in operation between June 1993 and June 1997. Pharmacies entered the demonstration in two waves, and were randomly asssigned to treatment and control groups in clusters. This study assessed the effects of the Iowa OPDUR demonstration on a set of adverse, clinical outcomes that have been linked to inappropriate prescribing of cardiovascular drugs. The population of interest in this study was elderly, Iowa residents who were eligible for Medicare and Medicaid, and who filled a prescription at a demonstration pharmacy at least once.

Rates of hyperkalemia, cardiac conditions, and symptoms and diagnoses of digoxin toxicity were expected to decline in the treatment group if the OPDUR system was effective in improving the prescribing of cardiovascular drugs. Holding other factors constant, treatment and control group patients generally did not have statistically significant differences in these outcome rates. The analysis presented in this chapter therefore indicates that the Iowa OPDUR system was *not* effective in preventing adverse outcomes associated with inappropriate prescribing of cardiovascular drugs.

### Adverse Clinical Outcomes Related to Use of Cardiovascular Drugs

This study focuses on the use of three cardiovascular drugs: ACE inhibitors, verapamil, and digoxin. Clinical research has indicated that the use of these three drugs is associated with a number of adverse outcomes, including hyperkalemia, cardiovascular problems, and digoxin toxicity. In the following sections, the relationships of interest in this study are motivated by a brief literature review of recent clinical research in this area.

### *ACE Inhibitors and Hyperkalemia*

The relationship between use of ACE inhibitors and hyperkalemia, a condition that is characterized by excessive amounts of potassium in the blood, is of particular interest in this study. Among patients with renal insufficiency, the body's potassium balance is maintained through an increase in excretion in the remaining nephrons of the kidney. This compensatory action is mediated by an increase in aldosterone. ACE inhibitors can suppress the alodosterone produced by the adrenal gland, causing potassium to accumulate [Schlueter et. al. (1994)]. Potassium can also accumulate if ACE inhibitors are prescribed with potassium or potassium sparing diuretics.

Hyperkalemia is is more common in patients with chronic renal failure, diabetes, autonomic neuropathy, and adrenal insufficiency [Howes (1995)]. Risk groups for hyperkalemia also include users of certain drugs, including salt substitutes, some potassium sparing diuretics, ACE inhibitors, NSAIDS, beta-adrenergic antagonists, cyclosporine, heparin, and digoxin. Although patients with advanced renal failure frequently have mild hyperkalemia, life threatening hyperkalemia is uncommon until the glomerular filtration rate (GFR) falls below 5-10 mL per minute. In almost all cases, hyperkalemia can be controlled by taking patients off the offending drugs and/or by administering other medications such as cation exchange resin, insulin and glucose [Rakel (1996)].

In a hospital study conducted through the Vermont College of Medicine, 308 patients (1.7 percent of all patients admitted over the course of one year) were identified as having severe hyperkalemia, and out of 172 were found to have sustained hyperkalemia. The authors reported that 43 (14.0 percent) of those patients sustained cases were found to be temporally related to a drug or drugs that could cause hyperkalemia. The most common drug associated with hyperkalemia in this population was potassium chloride, but catopril was implicated in about 20 percent of cases. Many of the patients with severe hyperkalemia had predisposing conditions, mostly renal insufficiency and diabetes [Rimmer et. al. (1987)].

The results of this study do not provide a clear sense of either incidence or prevalence, and they cannot be generalized to settings outside the hospital. Nevertheless, the results do document that hyperkalemia is linked to use of ACE inhibitors with some frequency, especially in certain high-risk groups.

### *Verapamil and Cardiovascular Events*

Verapamil is a calcium channel blocker used to manage hypertension and to prevent and control angina pectoris. The action of verapamil is as a vasodilator; it also acts as a depressant to myocardial contractility and the caridiac conduction system, particularly at the AV node. Calcium channel blockers as a class have been shown to be effective in lowering blood pressure, in relieving symptoms in patients with stable angina, and in increasing excerise tolerance, but not necessarily in preventing serious morbidity and mortality [Silvestry and Kimmel (1996)].

There has been concern that verapamil may increase the risk of adverse, cardiovascular outcomes, particularly when used with beta blockers or with other calcium channel blockers, or when used at certain dosage levels. A number of studies have focused on morbidity and mortality associated with verapamil following an acute myocardial infarction. The Danish Verapamil Infarction Trials (DAVIT) I and II were randomized, double blind, placebo-controlled trials. There were 717 treatment group patients (who received verapamil) and 719 controls (who received placebo) in the first trial, and 878 treatment group patients and 897 controls in the second trial [The Danish Study Group on Verapamil in Myocardial Infarction (1990)].

In the first DAVIT trial, reinfarction and mortality rates were lower in the treatment group than the control group, but this difference was not statistically significant. The second DAVIT trial showed another non-significant difference between the treatment and control groups in mortality rates, but there was a statistically significant reduction in first major event rates at 18 months in the treatment group. Sub-group analyses indicated that treatment group patients without heart failure had statistically significant reductions in mortality, cardiac events, and renal infarction compared to control group patients without heart failure. In a sub-group of people *with* existing heart failure, there were no statistically significant differences between the treatment and control groups.

When verapamil is used in combination with a beta blocker, it is possible that the inotropic, chronotropic, and dromotropic effects of the two drugs will be additive. This effect could result in am increased probability of conduction defects and heart failure [Sever (1989)]. Opie (1988), in his description of the addition of verapamil to a regimen that already includes propranolol or metoprolol, cites a number of studies that inidcate that this combination may lead to adverse cardiovascular outcomes. These adverse cardiovascular outcomes include reduced contractility [Packer et. al. (1982)], increased heart size [Johnstone et. al. (1985)], symptomatic sinus bradycardia [Winniford et. al. (1985)], and depressed left ventricular function [Johnstone et. al. (1985)]. Peargien and Benowitz (1991) note that when verapamil is given in combination with beta blockers, bradycardia, hypotension, and conduction disturbances may occur, especially in the presence of underlying heart disease [Hamann et. al. (1985, 1987)].

Researchers have found that verapamil use is associated with hypotension. Pearigen and Benowitz (1991) report that out of a series of 91 patients who were identified through a regional poison control center for overdose of calcium channel blockers, nine out of the 38 patients were hypotensive [Ramoska et. al. (1990)]. In a review of drugs that can cause orthostatic hypotension, Schoenberger (1991) classifies calcium channel blockers as infrequent in provoking this outcome, while dopamine agonists, for example, were considered to cause this outcome frequently, and for tricyclic antidepressants, hypotension was a common problem. Schoenberger (1991) cites MacMahon (1990) as reporting that hypotension may occur in approximately five percent of patients taking dihydropyridine calcium channel blockers, and in less than 2.5 percent of patients in the other classes of calcium channel blockers.

### *Digoxin and Digoxin Toxicity*

Digoxin has been a major component of therapy for controlling heart rate during atrial fibrilation. Digoxin has a narrow therapeutic window, however, and toxicity is a risk. Furthermore, digoxin can interact with other drugs that may be used to treat atrial fibrilation, such as verapamil, quinidine, and procainamide. Interaction with verapamil may reduce clearance of digoxin, resulting in an increase of blood digoxin levels of up to 60-90 percent [Opie (1988); Opie citing Pedersen (1985)]. Digoxin toxicity is most common among the elderly.

### Evaluation Approach

The objective of this analysis was to explore the relationship between the Iowa OPDUR system and the rate of adverse events related to use of cardiovascular drugs. It was hypothesized that pharmacists using the OPDUR system would be more likely than pharmacists not using the OPDUR system to detect prescription problems related to cardiovascular drug use and, consequently, to have an opportunity to prevent adverse outcomes. Patients who visited pharmacies where prescriptions were screened by OPDUR, therefore, should be less likely than control group patients to experience illnesses related to inappropriate prescribing of cardiovascular drugs.

The general objective of this study was to compare rates of clinical outcomes among patients who went to intervention pharmacies to those among patients who went to control pharmacies. Multiple regression was used in order to control for demographic, disease, and drug factors predisposing a given patient to the outcome under study. All control variables were measured during a baseline period before the demonstration began so as to be independent of any demonstration effects. A version of White's method that accounts for cluster based sampling was used to estimate standard errors [Rogers (1993)].

### Data

The data used in this study was extracted from Medicaid drug claims and Medicare health services utilization claims for patients in the state of Iowa in 1993 (to provide a baseline) and 1994-1996. (OPDUR was in operation in Iowa between June 1993 and June 1997.) Other records and data streams generated by the demonstration were also incorporated into the data set. Data development is described in detail in Chapter 5.7.

### Study Populations

As displayed below in Exhibit 5.11.1, the data was used to create four study populations: (1) a main study population; (2) a sub-sample of the main study population consisting of those who had been using ACE inhibitors, calcium channel blockers, and/or digoxin during the baseline period; (3) a sub-sample of the main study population consisting of those who had been using ACE inhibitors, calcium channel blockers, and/or digoxin during the baseline or the follow-up period.; and (4) a sub-sample of the main study population whose prescriptions triggered screen failures related to the use of cardiovascular drugs. Although all models were implemented using these four samples, not all results are presented in this chapter. Instead, this chapter focuses on interpreting and summarizing the results from these models; selected models estimated using the main population are presented in full in the appendix to this chapter.

**Exhibit 5.11.1**

**HCFA DUR Demonstration Evaluation**
**Study Populations**

| Study Population: | Description: |
|---|---|
| Main Population<br>(N = 4220) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies |
| Baseline and Follow-up Cardiac Drug Users<br>(N = 2515) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled at least one prescription for ACE inhibitors, calcium channel blockers, and/or digoxin during the baseline period or during the follow-up period |
| Baseline Cardiac Drug Users<br>(N = 2062) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, filled at least one prescription for ACE inhibitors, calcium channel blockers, and/or digoxin during the baseline period |
| Triggered a Cardiovascular Drug Screen Failure<br>(N = 1168) | Dually eligible for Medicare and Medicaid, at least 65 years old, filled a prescription at an active demonstration pharmacy, did not fill prescriptions at both intervention and control pharmacies, triggered at least one screen failure related to cardiovascular drug use |

Sources: IA Medicaid claims

The main study population consisted of elderly patients (65 and over on January 1, 1993) who were eligible for both the Medicaid and Medicare programs and who went to either intervention or control pharmacies during the intervention. The elderly are particularly at risk for adverse outcomes related to cardiovascular drug use. Moreover, they are generally in more delicate health and taking more drugs than younger patients; the elderly thus are hypothesized to be more likely than other people of other age groups to benefit from the intervention. The requirement of dual eligibility ensures that the analyses have complete and consistent records of drug and health services utilization for all patients in the data set.

The study population was further limited to individuals who used cardiac drugs during the baseline period, and individuals who used cardiac drugs during the baseline or follow-up periods. Narrowing the population in this way was intended to improve the specificity of the analysis.[39] Furthermore, the analysis considered both a sub-sample of baseline cardiac drug users as well as a sub-sample of baseline or follow-up cardiac drug users because of concerns that the intervention might affect cardiac drug use in the follow-up period. In this case, estimates based on the baseline/follow-up cardiac drug sub-sample might be biased.[40]

The models also were estimated using a sub-sample that consisted of those patients whose prescriptions had triggered at least one screen failure. In the treatment group, the pharmacist received the screen failure information on his/her computer and presumably read and reacted to the information.[41] In the control group, screen failure information was not transmitted to the pharmacist. Instead, in the control group, the screen failure information was recorded for research purposes only. In the sub-population of persons who triggered screen failures, therefore, this analysis compares the rate of adverse outcomes between indivduals whose pharmacists were alerted via OPDUR to a cardiovascular drug prescription problem and individuals whose pharmacists were not alerted via OPDUR to a cardiovascular drug prescription problem. The screen trigger sample is the most targeted in the sense that every individual in the sample had a cardiovascular drug prescribing problem; compared to the other samples analyzed, this group was at highest risk of adverse effects. Nevertheless, the sample size is reduced considerably by excluding those who did not trigger a screen failure.

### *Definitions of Baseline and Follow-up Periods*

Pharmacies enrolled in the Iowa OPDUR demonstration in two waves (June 1994 and December 1994). Although other approaches would be possible, this study assumes that a patient's exposure to the demonstration (and consequently the patient's follow-up period) began when a patient first filled a prescription at a pharmacy participating in the demonstration. In the majority of cases, the follow-up period then continued until the last month of 1996.[42] In certain cases, if the patient lost eligibility for Medicare or Medicaid, entered a nursing home, or entered an HMO, the follow-up period ended sooner because the outcomes data captured in the claims would be incomplete after that event.[43] A patient's death was also deemed to terminate the follow-up period.

In a similar way, patients' baseline periods differed in length. The baseline period was determined to end the day before the patient's first contact with a demonstration pharmacy and to reach back twelve months or until the patient first became eligible for the Medicare program, became eligible for the

---

[39] Nevertheless, analyses based on this more narrow population would miss any intervention effect that involved patients avoiding the use of cardiovascular drugs altogether.

[40] If the Iowa OPDUR demonstration affected cardiac drug use in the follow-up period and the population of interest consists of cardiac drug users in the baseline and follow-up periods, a selection problem would exist. In this case, it's possible that those who benefitted from the intervention would be more likely than other to be excluded from the sample.

[41] This reaction, of course, may have been a decision to do nothing.

[42] The OPDUR system was in operation through June 1997. Complete data, however, was available only through 1996.

[43] Nursing home prescriptions that were batched and sent to demonstration pharmacies were not screened using OPDUR.

Medicaid program, left an HMO, or left a nursing home and became resident in the community. Again, the latter three restrictions are imposed in order to ensure that the medical record is complete during the baseline period.

Note that clinical outcomes that occurred on the day the patient first went to a demonstration pharmacy are excluded from both the baseline and the follow-up periods. In such cases, it is impossible to know whether the clinical issue precipitated the prescription or the prescription led to the clinical issue. The safe decision therefore was to omit the outcome from both periods.

### *Dependent Variables*

Based on the research literature, reviewed briefly above, three outcomes were selected for this study: (1) hyperkalemia, related to the use of ACE inhibitors; (2) congestive heart failure and other adverse cardiac events related to the use of verapamil, a calcium channel blocker; and (3) diagnosis and symptoms of digoxin toxicity, related to the use of digoxin. To capture cardiac problems related to verapamil use, two measures were constructed; one measure included diagnosis of congestive heart failure only, while the other, broader measure included congestive heart failure, conduction disorders, sinus bradycardia and hypotension. Digoxin toxicity was measured using four different variables that ranged from a specific diagnosis of digoxin toxicity to a broad range of symptoms that are associated with digoxin toxicity. These dependent variables are summarized below in Exhibit 5.11.2.

Outcome indicators were constructed using Medicare diagnosis codes.[44] Each outcome was associated with a series of ICD-9 codes. If a patient's record showed evidence of one of those ICD-9s as either the primary or secondary diagnosis for an inpatient stay or for a physician encounter (paid by Part A or Part B of Medicare) during the follow-up period, then the patient was determined to have had the outcome.

For each outcome, a model also was estimated in which the outcome variable was based on inpatient claims alone. It is possible that compared to physicians in the outpatient setting, hospitals might be more precise in their coding of diagnoses because of its payment implications. Variables that were created using inpatient claims only therefore might contain less measurement error compared to variables created using all claims. Models based on variables created using inpatient claims only are not presented here because the results were very similar to models based on variables created using all claims.

### *Independent Variables*

Each model included a basic set of covariates that captured exposure to the intervention as well as general predisposing and need characteristics. Additionally, covariates were added to each model to control for clinical risk factors that are specific to each of the outcomes. The main set of independent

---

[44] Because all of the samples used in this analysis included individuals who were dually eligible for Medicare and Medicaid, it is assumed that all medical services were initially billed and primarily covered by Medicare.

**Exhibit 5.11.2**

**HCFA DUR Demonstration Evaluation Dependent Variables**

| Variable: | ICD-9 Classification: |
|---|---|
| hyperkalemia | 276.7 |
| congestive heart failure (CHF) | 428.0 |
| congestive heart failure or conduction disorders or hypotension or sinus bradycardia (CHF+) | 428.0, all 426, 458.0-458.9, 427.81-427.89 |
| any one of digoxin toxicity diagnosis, atrial ventricular block, conduction disorders, sinus bradycardia, ventricular tachycardia, abdominal pain, anorexia, disturbed vision, fatigue, headache or nausea/ vomiting (all symptoms and diagnosis) | 972.1, 426.00-426.13, all 426, 427.81-427.89, 427.1, 789.00-789.09, 783.0, 368.9, 780.7, 784.0, 787.01-787.03 |
| digoxin toxicity diagnosis only | 972.1 |
| arrhythmia complex only (any one of atrial ventricular block, conduction disorders, sinus bradycardia, or ventricular tachycardia) | 426.00-426.13, all 426, 427.81 -427.89, 427.1 |
| symptom complex only (any one of abdominal pain, anorexia, disturbed vision, fatigue, headache, or nausea/vomiting) | 789.00-789.09, 783.0, 368.9, 780.7, 784.0, 787.01-787.03 |

*Sources:* International Classification of Diseases, 9th Revision, DHHS Publication No. (PHS) 91-1260, October, 1991.

variables, as well as the outcome specific independent variables, are described in the following sections.

***Exposure to the Intervention.*** Although this demonstration was designed as a random experiment, randomization occurred at the level of the pharmacy cluster while clinical benefits accrued at the level of the patient.[45] This posed several methodological challenges. First, individuals had the potential to go to multiple pharmacies and, potentially, to go to a mix of intervention, control, and non-demonstration pharmacies. The study adopted a simple approach. Individuals were considered members of the intervention group if they ever had a Medicaid-paid drug claim at an intervention pharmacy and never had such a claim at a control pharmacy. They were considered members of the control group if they ever had a Medicaid-paid drug claim at a control pharmacy and never had such a claim at an intervention pharmacy. Individuals who went to both intervention and control pharmacies were excluded from the study.[46]

---

[45] Pharmacy clusters, built around common prescribers, were created and randomized to minimize the influence that intervention pharmacies might exert through common prescribers on control patients. In IA, there were 191 pharmacy clusters ranging in size from 1 person to 144 people.

*General Predisposing and Need Characteristics.* The basic design of the study was to compare treatment and control patients within a regression framework that controls for demographic, drug, and disease factors. For all five outcomes, regressions included controls for the patient's race, sex, and age. These factors affect health outcomes both directly and as proxies for behavior; they may also affect the predisposition to use care. These variables, which were taken from the enrollment records of the Medicaid program, were included in regressions as a series of indicator variables.

### *Length of Baseline and Follow-up Periods*

Other things equal, a patient with a longer follow-up period would be more likely to experience any clinical outcome than one with a shorter follow-up period, thus the length of the follow-up period was made an additional control in models. While the length of the baseline period does not affect patients' probabilities of clinical problems in and of itself, expected health spending, number of drugs, and risk factors will tend to be slightly under measured for patients with short baselines relative to patients with full baselines. Thus, controlling for measured covariates, the probability of clinical outcomes might be higher for patients with short baselines; consequently, the length of the baseline period was included as an additional control. The lengths of the follow-up and baseline period were entered in models as the natural log of the number of months in each period.

### *Pharmacy Wave*

Pharmacies entered the cognitive services demonstration in two waves. The first wave of 119 pharmacies enrolled in June 1994. A second wave of 104 pharmacies entered the demonstration in December 1994. It is possible that pharmacies entering the demonstration later learned from the experiences of prior entrants and used the OPDUR system more effectively. Indicator variables were included in the models that capture the wave of the first demonstration pharmacy where the individual filled a prescription. The omitted category in the models was wave one.

The basic set of independent variables that were used in every model (as described above) are summarized below in Exhibit 5.11.3.

## Risk Factors Specific to Each Outcome

*Hyperkalemia.* Analyses of hyperkalemia added additional controls for pre-existing hyperkalemia, baseline renal insufficiency, baseline congestive heart failure, and baseline diabetes, as captured by baseline claims for insulin and oral hypoglacemics. These conditions are risk factors for ongoing conduction disorders and cardiac dysrhythmia. The models also controlled for the use of beta blockers in the baseline period because use of these drugs is associated with hyperkalemia.

---

[46] For this purpose, the day that the patient first went to the demonstration pharmacy was included in the follow-up period.

**Exhibit 5.11.3**

**HCFA DUR Demonstration Evaluation**
**Basic Set of Covariates Included in All Models**

race
sex
age (in categories)
length of baseline period
length of follow-up period
wave of first contact demonstration pharmacy

***Congestive Heart Failure/Conduction Disorders/Hypotension/Sinus Bradycardia.*** Analyses of these outcomes controlled for pre-existing congestive heart failure, baseline conduction disorders, baseline hypotension, baseline sinus bradycardia, and other cardiac conditions diagnosed during the baseline period.

***Digoxin Toxicity Diagnosis/Symptom Complex/Arrythmia Complex.*** Analyses of these outcomes controlled for baseline digoxin toxicity diagnosis and symptoms, baseline renal insufficiency, baseline congestive heart failure, and baseline diabetes (as captured by claims for insulin and oral hypoglacemics). These conditions are risk factors for digoxin toxicity. The use of certain drugs in combination with digoxin is also associated with digoxin toxicity. For this reason, baseline use of verapamil, quinidine, amiodarone, propafenone, indomethacin, itraconazole, aplrazolam, and diuretics were included as covariates in the digoxin toxicity models.

## Rate of Clinical Outcome in Pharmacy Cluster

The pharmacy cluster, the unit of randomization in this analysis, effectively defined medical communities. Potentially, there could be unmeasured differences in demographic and behavioral characteristics among patients in different pharmacy clusters that influenced the need for care and the predisposition to seek it. Similarly, there could be geographic variation associated with cluster in access to care and in its practice, such as variations in coding of conditions and diagnoses and in treatment protocols. If these differences were associated with measured clinical outcomes, then controlling for them would improve the precision of estimated intervention effects. Thus, each model included the average outcome rate during the baseline period for each pharmacy cluster as an additional regressor. This variable served as a proxy for the net effect of geographic factors on the outcome of interest.

***Independent Variable Construction.*** The indicators for patients' baseline disease status were created in the same way as the indicators for patients' clinical outcomes. The appendix to this chapter shows the relevant ICD-9 codes. The indicators for patients' drug utilization were based on Medicaid drug claims during the baseline period. For all drugs, the indicator was positive based on a single prescription claim.

## Models and Statistical Methods

The fact that pharmacy clusters rather than individual patients were randomized also meant that the standard regression framework required some adjustment. If indeed there were factors correlated with cluster that affected outcomes and if, despite the controls included in the models, these factors were not fully measured, then, because the number of clusters was fairly small (191) relative to the number of patients, these differences might not average out between intervention and control groups, despite the large number of patients, leading to misleading estimates of intervention effects. Specifically, traditionally-estimated standard errors would be smaller than the true standard errors.

As has been discussed, the first line of defense against this problem was the effort to include a full set of covariates, in particular the cluster averages discussed above. In addition, a cluster adjusted version of White's method was used to estimate the variance-covariance matrices of the models. This method relaxes the assumption of independence within clusters, and is considered to be a robust variance estimator in cases of cluster-based sampling.[47] An alternative approach would have been to bootstrap the coefficients of each model and derive nonparametric standard errors for the coefficients.[48] For the main population, this alternative method was tested; each anti-depressant model was bootstrapped and standard errors were computed. The results, which are not presented here, were very similar to the results derived using the cluster-adjusted version of White's method.[49]

Thus, the regression specification was:

$$y_{ijt_f}^{c}=\alpha+\beta_1 tr_{ij}+X_{it_b}\beta_2+W_{it_b}^{c}\beta_3+Z_i\beta_4+\beta_5 y_{jt_b}^{c}+\nu_{ij}$$

with:

| | |
|---|---|
| $y^c_{ijtf}$ | indicator that patient i associated with pharmacy cluster j experienced clinical outcome c during the follow-up period (tf) |
| $tr_{ij}$ | indicator that patient i filled a prescription at treatment pharmacies and not at control pharmacies during the demonstration period |
| $X_{it}$ | general predisposing (socio-demographic) and need characteristics of patient i determined during the baseline period |

---

[47] William Rogers originally developed this method in the early 1980's to account for cluster-based sampling issues in the RAND Health Insurance Experiment.

[48] The bootstrapping method involves pulling repeated samples (with replacement), estimating coefficients repeatedly (100 times in this case), and then deriving standard errors from the distributions of estimated coefficients. This method is often used in cases where we have no formulas to compute standard errors or when the assumptions we need to make to compute standard errors do not make sense (i.e. in this case, we could not assume that error terms were independently distributed because of inter-cluster correlation).

[49] Two other methods were considered to account for the effects of clustered sampling. The random effects probit was considered, but the large size of some of the clusters made this model very difficult if not impossible to implement computationally. Generalized Estimating Equations (GEE) methods could have been used, but this approach would have made estimation more complex and the advantages of this approach were uncertain.

| | |
|---|---|
| $W^{c}_{itb}$ | patient i's risk factors for outcome c determined during the baseline period |
| $Z_i$ | length of follow-up and baseline windows for patient i |
| $y^{c}_{jtb}$ | rate of clinical outcome c in pharmacy cluster j during the baseline period |
| $\nu_{ij}$ | error |

Because all dependent variables were binary, probit regressions were used. The probit coefficients were transformed to marginal effects to make interpretation easy.[50]

Exhibit 5.11.4 below summarizes all of the models that were estimated in this study. For all of the outcomes except new hip fractures, models were estimated using variables created from all Medicare claims and variables created from inpatient Medicare claims only. Each model was estimated using the four populations described previously: (1) the main population of dually eligible beneficiaries; (2) a sub-population of dually eligible beneficiaries who used cardiac drugs in the baseline or follow-up periods; (3) a sub-population of dually eligible beneficiaries who used cardiac drugs in the baseline period only; and (4) a sub-population of dually eligible beneficiaries who triggered at least one cardiovascular drug screen failure. Only models based on all claims and the main population are presented in full in Appendix I (these models are shaded in Exhibit 5.11.4).

## Findings

### *Descriptive Statistics*

Exhibit 5.11.5 displays the means and standard deviations for the dependent variables and the basic set of covariates used in all of the models described above. The treatment group had a larger proportion of non-white individuals relative to the comparison group. Additionally, the treatment group had a marginally significant, lower rate of digoxin toxicity symptoms, and a marginally significant, higher rate of hyperkalemia compared to the comparison group. Overall, the descriptive statistics reveal no striking evidence of a treatment effect;the treatment and comparison groups were quite similar along both baseline characteristics and follow-up outcomes.

---

[50] The marginal effect can be interpreted as the change in the probability of the outcome that is associated with an infinitesimal change in the independent variable, holding other covariates constant. For binary independent variables, the marginal effects are calculated based on a one unit change in the independent variable rather than an infinitesimal change.

Exhibit 5.11.4

## HCFA DUR Demonstration Evaluation
## Models Estimated

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 1 | Hyperkalemia | Probit | All dual eligibles | 4220 | Based on all Medicare claims. |
| 2 | Hyperkalemia | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |
| 3 | Hyperkalemia | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on all Medicare claims. |
| 4 | Hyperkalemia | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on inpatient Medicare claims only. |
| 5 | Hyperkalemia | Probit | Cardiac drug users in baseline | 2062 | Based on all Medicare claims. |
| 6 | Hyperkalemia | Probit | Cardiac drug users in baseline | 2062 | Based on inpatient Medicare claims only. |
| 7 | Hyperkalemia | Probit | Those who triggered cardiovascular drug screen failures | 1168 | Based on all Medicare claims. |
| 8 | Hyperkalemia | Probit | Those who triggered cardiovascular drug screen failures | 1168 | Based on inpatient Medcare claims only. |
| 9 | Congestive Heart Failure | Probit | All dual eligibles | 4220 | Based on all Medicare claims. |
| 10 | Congestive Heart Failure | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |
| 11 | Congestive Heart Falure | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on all Medicare claims. |
| 12 | Congestive Heart Failure | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on inpatient Medicare claims only. |
| 13 | Congestive Heart Failure | Probit | Cardiac drug users in baseline | 2062 | Based on all Medicare claims. |
| 14 | Congestive Heart Failure | Probit | Cardiac drug users in baseline | 2062 | Based on inpatient Medicare claims only. |
| 15 | Congestive Heart Failure | Probit | Those who triggered cardiovascular drug screen faillure | 1168 | Based on all Medicare claims. |
| 16 | Congestive Heart Failure | Probit | Those who triggered cardiovascular drug screen faillure | 1168 | Based on inpatient Medicare claims only. |
| 17 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | All dual eligibles | 4220 | Based on all Medicare claims. |

Exhibit 5.11.4

## HCFA DUR Demonstration Evaluation
## Models Estimated

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 18 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |
| 19 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on all Medicare claims. |
| 20 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on inpatient Medicare claims only. |
| 21 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | Cardiac drug users in baseline | 2062 | Based on all Medicare claims. |
| 22 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | Cardiac drug users in baseline | 2062 | Based on inpatient Medicare claims only. |
| 23 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | Those who triggered cardiovascular drug screen failure | 1168 | Based on all Medicare claims. |
| 24 | Congestive Heart Failure, Conduction Disorders, and/or Sinus Bradycardia | Probit | Those who triggered cardiovascular drug screen failure | 1168 | Based on inpatient Medicare claims only. |
| 25 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | All dual eligibles | 4220 | Based on all Medicare claims. |
| 26 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |
| 27 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on inpatient Medicare claims only. |
| 28 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on all Medicare claims. |
| 29 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | Cardiac drug users in baseline | 2062 | Based on all Medicare claims. |

**Exhibit 5.11.4**

**HCFA DUR Demonstration Evaluation**
**Models Estimated**

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 30 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | Cardiac drug users in baseline | 2062 | Based on inpatient Medicare claims only. |
| 31 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | Those who triggered cardiovascular drug screen failure | 1168 | Based on all Medicare claims. |
| 32 | Digoxin Toxicity Diagnosis and/or Any Symptoms | Probit | Those who triggered cardiovascular drug screen failure | 1168 | Based on inpatient Medicare claims only. |
| 33 | Digoxin Toxicity Diagnosis | Probit | All dual eligibles | 4220 | Based on all Medicare claims. |
| 34 | Digoxin Toxicity Diagnosis | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |
| 35 | Digoxin Toxicity Diagnosis | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on all Medicare claims. |
| 36 | Digoxin Toxicity Diagnosis | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on inpatient Medicare claims only. |
| 37 | Digoxin Toxicity Diagnosis | Probit | Cardiac drug users in baseline | 2062 | Based on all Medicare claims. |
| 38 | Digoxin Toxicity Diagnosis | Probit | Cardiac drug users in baseline | 2062 | Based on all Medicare claims. |
| 39 | Digoxin Toxicity Diagnosis | Probit | Those who triggered cardiovascular drug screen failure | 1168 | Based on all Medicare claims. |
| 40 | Digoxin Toxicity Diagnosis | Probit | Those who triggered cardiovascular drug screen failure | 1168 | Based on all Medicare claims. |
| 41 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | All dual eligibles | 4220 | Based on all Medicare claims. |
| 42 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |
| 43 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on all Medicare claims. |
| 44 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on inpatient Medicare claims only. |

Exhibit 5.11.4

## HCFA DUR Demonstration Evaluation
## Models Estimated

| Model Number: | Dependent Variable: | Estimation Method: | Population: | N | Variable Construction: |
|---|---|---|---|---|---|
| 45 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | Cardiac drug users in baseline | 2062 | Based on all Medicare claims. |
| 46 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | Cardiac drug users in baseline | 2062 | Based on inpatient Medicare claims only. |
| 47 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | Those who triggered screen failure | 1168 | Based on all Medicare claims. |
| 48 | Digoxin Toxicity - Arrhythmia Complex Symptoms Only | Probit | Those who triggered screen failure | 1168 | Based on inpatient Medicare claims only. |
| 49 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | All dual eligibles | 4220 | Based on all Medicare claims. |
| 50 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | All dual eligibles | 4220 | Based on inpatient Medicare claims only. |
| 51 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on all Medicare claims. |
| 52 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | Cardiac drug users in baseline or follow-up | 2515 | Based on inpatient Medicare claims only. |
| 53 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | Cardiac drug users in baseline | 2062 | Based on all Medicare claims. |
| 504 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | Cardiac drug users in baseline | 2062 | Based on inpatient Medicare claims only. |
| 55 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | Triggered cardiovascular drug screen failure | 1168 | Based on all Medicare claims. |
| 56 | Digoxin Toxicity - Symptom Complex Symptoms Only | Probit | Triggered cardiovascular drug screen failure | 1168 | Based on inpatient Medicare claims only. |

*Sources:* International Classification of Diseases, 9th Revision, DHHS Publication No. (PHS) 91-1260, October, 1991.

**Exhibit 5.11.5**

### HCFA DUR Demonstration Evaluation
### Means and Standard Deviations of Dependent Variables and Covariates in Main Iowa Population Models
### (N = 4220)

| | Treatment<br>(N = 1974) | Control<br>(N = 2246) |
|---|---|---|
| Dependent Variables | | |
| hyperkalemia* | .031<br>(.004) | .022<br>(.003) |
| congestive heart failure | .231<br>(.009) | .252<br>(.009) |
| congestive heart failure, conduction disorders, hypotension, or sinus bradycardia | .307<br>(.010) | .321<br>(.010) |
| digoxin toxicity diagnosis and/or any symptoms* | .384<br>(.011) | .416<br>(.010) |
| digoxin toxicity diagnosis | .003<br>(.001) | .003<br>(.001) |
| digoxin toxicity arrhythmia complex only | .120<br>(.007) | .114<br>(.007) |
| digoxin toxicity symptom complex only* | .332<br>(.011) | .364<br>(.010) |
| Main Covariates | | |
| female | .820<br>(.009) | .823<br>(.008) |
| white*** | .866<br>(.008) | .934<br>(.005) |
| black*** | .057<br>(.005) | .034<br>(.004) |
| other race*** | .076<br>(.006) | .032<br>(.004) |
| age 65-69 | .213<br>(.009) | .204<br>(.009) |
| age 70-74 | .210<br>(.009) | .210<br>(.009) |
| age 75-79 | .160<br>(.008) | .178<br>(.008) |
| age 80-84 | .133<br>(.008) | .134<br>(.007) |
| age 85 or higher | .121<br>(.007) | .110<br>(.007) |
| length of baseline period* | 10.90<br>(.059) | 11.02<br>(.053) |

**Exhibit 5.11.5**

**HCFA DUR Demonstration Evaluation**
**Means and Standard Deviations of Dependent Variables and Covariates in Main Iowa Population Models**
**(N = 4220)**

| | | |
|---|---|---|
| length of follow-up period | 17.54<br>(.213) | 17.90<br>(.233) |

Note: *** statistically significant difference at the .01 level
** statistically significant difference at the .05 level
* statistically significant difference at the .10 level

## Results

### *Treatment Effects*

It was hypothesized that patients who visited pharmacies that were receiving OPDUR messages would be less likely than similar patients visiting other pharmacies to experience adverse health outcomes related to improper prescribing of cardiovascular drugs. This analysis focused on three, preventable adverse outcomes related to specific types of cardiovascular drug use: (1) ACE inhibitors and hyperkalemia; (2) verapamil, a calcium channel blocker, and cardiac problems; and (3) digoxin and digoxin toxicity. If the OPDUR messages did significantly decrease rates of these problems, the models should indicate that, after adjusting for baseline factors, the treatment group had lower rates of adverse outcomes relative to the control group.

There was no evidence that the intervention was effective in reducing rates of hyperkalemia. In these models, the treatment effect was never statistically significant at the .05 level. Similarly, the treatment effect was never statistically significant at the .05 level in any of the congestive heart failure models. These findings indicate that the OPDUR system was ineffective in preventing both hyperkalemia related to inappropriate prescribing of ACE inhibitors, and congestive heart failure and other serious cardiac conditions related to the improper use of verapamil, a calcium channel blocker.

The effects of the intervention on rates of digoxin toxicity were assessed by using a variety of outcome measures, ranging from a diagnosis of digoxin toxicity to a broad range of symptoms that are associated with digoxin toxicity. It was expected that diagnoses of digoxin toxicity would be rare, while the symptoms of digoxin toxicity (which include common ailments such as nausea and headache) would be quite prevalent. If the models indicated that treatment group patients were less likely to experience both the symptoms of digoxin toxicity and a diagnosis of digoxin toxicity, ths analysis would suggest that the OPDUR system was effective in preventing digoxin toxicity.

Across all of the models where symptom complex (which consisted of abdominal pain, anorexia, disturbed vision, fatigue, headache and/or nausea/vomiting) was the dependent variable or part of the dependent variable, the treatment effect was negative and statistically significant at the .05 level. The magnitude of the decline in probability, which ranged from .04 to .07, was larger among subpopulations of cardiovascular drug users and the sub-population of individuals who triggered cardiovascular drug screen failures. These results offer suggestive evidence that the Iowa OPDUR system was associated with a decrease in digoxin toxicity, and that the system's effectiveness was magnified slightly for high risk sub-populations.

If the Iowa OPDUR system did reduce the risk of digoxin toxicity, however, one would expect that the intervention would be associated with both a decline in diagnoses of digoxin toxicity and a decline in cardiac symptoms related to the disease. Surprisingly, the models displayed no evidence of this kind. The treatment effect generally was not statistically significant in models where digoxin toxicity diagnosis only, or arrythmia complex symptoms only, were dependent variables. These findings contradict the previous result, which suggested that the OPDUR system may have been effective in reducing the risk of digoxin toxicity.

Overall, then, the digoxin toxicity models do not present strong evidence that the Iowa OPDUR system was effective in preventing digoxin toxicity. Although patients who visited pharmacies that received OPDUR messages were less likely than others to experience abdominal pain, anorexia and other common symptoms of digoxin toxicity, they were just as likely as control group patients to have cardiac symptoms of digoxin toxicity or a diagnosis of digoxin toxicity. It appears that as the outcome variable was made more narrow and specific in the models, the association between the intervention and digoxin toxicity disappeared. This finding suggests that a relationship between the OPDUR system and rates of digoxin toxicity probably does not exist. It remains unclear, therefore, why the intervention was associated with a small but statistically significant drop in the probability of commonly reported ailments that are symptoms of digoxin toxicity.

Exhibit 5.11.6 below displays the signs of the treatment coefficients across the seven outcome measures. Although the treatment coefficients were generally negative, the magnitudes of the coefficients were usually close to zero. The treatment coefficient was only statistically significant in the models where symptoms of digoxin toxicity were part of the dependent variable. The results do not indicate, therefore, that OPDUR was effective in preventing adverse outcomes related to cardiovascular drug use.

### Other Covariates

Gender and race generally were not statistically significant predictors of outcomes. Blacks, however, were slightly more likely than individuals of other races to have a diagnosis of digoxin toxicity. Although there was no consistent relationship between all of the age categories and the outcomes, the two oldest sub-groups (80-85 and 85+ years) were more likely than younger individuals to experience all outcomes.

Baseline measures of diseases and drug utilization that were expected to place individuals at higher risk for outcomes were included as covariates in the models. These covariates were good predictors of the outcomes. Hyperkalemia in the follow-up period was significantly associated with baseline hyperkalemia as well as baseline renal insufficiency, baseline congestive heart failure, and baseline diabetes, as measured by the use of insulin and oral hypoglacemics. Congestive heart failure, hypotension, conduction disorders and sinus bradycardia in the follow-up period were significantly

**Exhibit 5.11.6**

**HCFA DUR Demonstration Evaluation**
**Summary of Signs of Treatment Effects Coefficients**
**Probit Models Based on Variables Created Using All Medicare Claims**

| Population | Dependent Variable | | | | | | |
|---|---|---|---|---|---|---|---|
| | Hyperkalemia | CHF | CHF+ | Digoxin Toxicity Symptoms or Diagnosis | Digoxin Toxicity Diagnosis | Arrythmia Complex | Symptom Complex |
| Main | .114<br>(.089) | -.057<br>(.048) | -.042<br>(.044) | -.109***<br>(.042) | -.217<br>(.314) | -.007<br>(.056) | -.108**<br>(.043) |
| Baseline & Follow-Up Users | .100<br>(.101) | -.039<br>(.057) | -.020<br>(.055) | -.160***<br>(.054) | -.001<br>(.003) | -.039<br>(.067) | -.150**<br>(.055) |
| Baseline Users only | .165<br>(.109) | -.068<br>(.064) | -.048<br>(.061) | -.157***<br>(.060) | -.217<br>(.314) | -.037<br>(.075) | -.160**<br>(.061) |
| Triggered Screen Failure | .129<br>(.142) | -.029<br>(.083) | -.017<br>(.081) | -.184***<br>(.081) | -.199<br>(.328) | -.071<br>(.096) | -.190**<br>(.080) |

Note: *** statistically significant difference between the arms at the 1 percent level
** statistically significant difference between the arms at the 5 percent level
* statistically significant difference between the arms at the 10 percent level
*Sources:* Iowa Medicare claims

associated with baseline cardiac conditions. Digoxin toxicity was significantly associated with baseline digoxin toxicity, baseline congestive heart failure, baseline diuretics use, and baseline diabetes, as measured by use of insulin and oral hypoglycemics.

It was expected that the length of the baseline and follow-up periods might affect outcomes because individuals with longer baseline periods might have better information about baseline diseases and drug utilization, and because individuals with longer follow-up periods might have more time under observation. This hypothesis was confirmed for the follow-up length variable; baseline length, however, was not a significant predictor of outcomes.

The wave variables were created to capture learning effects and other variation in responses to OPDUR that might be correlated with the pharmacy's date of entrance into the demonstration. The models revealed, however, that there was no statistically significant association between pharmacy wave and adverse outcomes related to cardiovascular drug use.

### *Pharmacy Cluster Effects*

To account for the fact that clusters of pharmacies rather than individuals were randomized to the treatment and control groups, each model included as a covariate the average rate of the dependent variable in the individual's pharmacy cluster. It was expected that this variable would help control for unobservable, cluster level differences in factors that affect the outcomes of interest. These variables were not consistently statistically significant across the models, and the magnitude of the marginal effect of the cluster rates on the outcomes was usually very close to zero. Additionally, the Huber/White [White (1980)] variance estimator, which also was used to adjust for cluster-based sampling, led to standard errors that were very similar to those calculated using a conventional variance estimate. Overall, then, adjustment for cluster-based sampling did not have any impact on the interpretation of the findings.

### Comment

The objective of the Iowa OPDUR system was to alert pharmacists to prescription problems at the point of service so that adverse outcomes could be avoided. This study focused on the impact of the intervention on a set of clinical outcomes related to cardiovascular drug use: (1) hyperkalemia; (2) congestive heart failure, hypotension, sinus bradycardia, and conduction disorders; and (3) digoxin toxicity diagnosis and clusters of symptoms related to digoxin toxicity. The analysis controlled for treatment and comparison group characteristics that may have differed at baseline, as well as the cluster-based sampling design used in the demonstration.

The findings indicate that the OPDUR system had no measurable impact on the outcomes of interest; apparently, the OPDUR messages related to cardiovascular drug therapy issues did not prevent adverse outcomes that would have occurred had the system not been in operation. There are a number of reasons why this effect might not exist, or why the effect might be too small to detect using this type of analysis. For an OPDUR message to prevent an adverse outcome, a chain of events, potentially involving pharmacists, prescribers and patients, must have taken place. Assuming that the OPDUR system correctly identified a potentially serious drug therapy problem, the pharmacist must have responded to the warning in a manner that reduced the risk of an adverse outcome. The pharmacist's actions may have involved the prescriber and/or the patient.

It is more likely, however, that limitations in the study design clouded measurement of an existing treatment effect, particularly if the treatment effect was very small. Generally, a cluster sampling approach reduces the power of detecting differences between the treatment and control groups. In this case, however, the intracluster correlation is likely to be low, and many of the clusters were very small; cluster based sampling therefore does not seriously reduce the power of the study design.

More important is the fact that the treatment effect is expected to be very small — it is hypothesized that the intervention might lead to half a percentage point difference in outcomes between the treatment and control groups.[51] With the sample sizes used in the sub-population of baseline and follow-up period cardiac drug users, the study probably had less than ten percent power to detect a difference of this size. To capture such tiny differences between treatment and control groups, therefore, huge sample sizes would be needed.[52] Despite this limitation and other constraints of the study design, the findings offer strong evidence that the Iowa OPDUR demonstration did not lead to large declines in the risk of adverse outcomes related to cardiovascular drug use.

---

[51] This expectation is based on the literature which suggests that about one percent of prescriptions will be associated with serious adverse clinical outcomes.

## 5.12 The Impact of the Maryland ProDUR Intervention for β-Agonist Inhaler Use on Elderly Dual-Eligible Medicaid Recipients

In July 1994, the State of Maryland implemented a set of prospective DUR screens intended to improve prescribing of β-agonist metered dose inhalers (MDIs) for Medicaid recipients suffering from various chronic respiratory diseases such as asthma and COPD. Part of the Maryland ProDUR system, the screening procedure involves an on-line review of each new β-agonist MDI prescription before it is filled. If the prescription violates one or more of the predetermined screening criteria, the pharmacist is alerted to the potential problem. The system uses "soft" alerts, meaning that pharmacists can override the DUR message and fill the prescription as written without seeking permission either from a system employee or the prescribing physician.

This study was designed to evaluate the impact of the β-agonist screening program on the health status of elderly Maryland Medicaid recipients who suffer from chronic respiratory disease. The study employed a pre-post, cross-sectional design with two comparator state, Georgia and Iowa. The analysis used Medicaid and Medicare enrollment and claims data to select the study populations, assess exposure to the ProDUR intervention, develop patient risk profiles, and tabulate rates of adverse clinical outcomes for each study subject. Descriptive time plots and regression models were used to test the hypothesis that ProDUR screening reduces the incidence of adverse effects asociated with misuse of β-agonist inhalers. The study findings produced no evidence that the screening program was effective in this regard.

### The Maryland OPDUR Process

The 1990 Omnibus Budget Reconciliation Act (OBRA) mandated that states establish a drug use review (DUR) board for outpatient drugs in the Medicaid population. Maryland has had a retrospective DUR program since 1986 and the on-line prospective DUR (OPDUR) program in place since 1993.

---

[52] The problem of inadequate power was anticipated early in the planning stages of this study. In an effort to increase sample sizes, dually eligible individuals from Maryland and Georgia who used ACE inhibitors, calcium channel blockers, or digoxin between 1993 and 1996 were considered as groups that could be used to augment the sample size. Since Maryland had an OPDUR system in operation between 1993 and 1996, Maryland cardiovascular drug users could be added to the treatment group. Georgia, a state that never had an OPDUR system in place during this time period, was considered suitable as a proxy for the control group. The addition of Maryland and Georgia beneficiaries to this analysis would have had a sizeable impact on power; it was estimated that if dually eligible, cardiovascular drug users from these states were added to the Iowa sample, the study would have 80 percent power to detect differences between .5 and 2 percent, depending on the outcome.

Incorporating Maryland and Georgia into the analysis has a number of potentially serious limitations. If the Iowa analysis had led to suggestive results, however, the addition of Maryland and Georgia data to the analysis still would have been useful as a method of confirming the findings with larger sample sizes. The Iowa analysis in this chapter, however, did not offer any evidence that the OPDUR system was effective in reducing adverse outcomes related to cardiovascular drug use. Consequently, it was decided that the disadvantages of adding data from other states to the study outweighed the advantages of greater power.

OBRA '90 requires that states establish DUR boards as state organizations or through contractual arrangements with public or private organizations. The Maryland Department of Health and Mental Hygiene opted to contract with the University of Maryland's Center on Drugs and Public Policy (CDPP) for administrative and technical support for assembling and staffing the DUR board. The vendor selected to conduct Maryland's prospective DUR claims evaluation was First Health Services Corporation.

The DUR board performs the following roles:

- advises the state on matters pertaining to drug use review;
- approves criteria and standards applied in prospective and retrospective DUR;
- evaluates the criteria and standards to identify potentially inappropriate or medically unnecessary care provided by physicians and pharmacists to individuals receiving benefits under Medicaid;
- oversees operation of the prospective DUR program by ensuring the criteria are applied consistently with those used in retrospective DUR; and
- identifies education needs and develops educational plans to improve prescribing or dispensing practices and evaluates their effects.

It is the policy of the Maryland DUR Board that prospective DUR criteria must meet the following conditions: (1) the potential consequences of an exception to the criterion are severe enough that immediate action by the pharmacist is warranted,[53] (2) it is well documented that the probability of an adverse event is high if an exception results and the situation is not addressed; and (3) the criterion is logistically feasible to implement.

The DUR Board employs a two-phase process for DUR criteria development and approval: a provisional and final approval phase. During the provisional approval phase, the Center with input from the DUR Board and the Maryland Medicaid program, identifies specific candidate drug groups for consideration based on frequency of use, cost or suspected inappropriate use. The Center then reviews existing retrospective and prospective DUR contractors' criteria, drug compendia, and peer-reviewed literature. Based on this information the Center then suggests criteria for each drug in the class, utilizing the criteria elements specified in OBRA '90. These draft criteria are presented to the DUR Board for comment and consideration.

---

[53] Severity may include situations where either the clinical or economic impact (or both) is judged to be significant.

The final approval phase begins with distribution of the provisionally approved criteria to the community for comment. Provisional criteria are made available to pharmaceutical industry personnel through a Pharmaceutical Manufacturers Association representative who regularly attends DUR Board meetings. Representatives from the retrospective and prospective DUR contractors also regularly attend Board meetings and offer advice and suggestions on criteria feasibility and operationalization. Anyone may request copies of the provisional criteria from the Center and may submit written comment on the criteria back to the Center. The Center then drafts final criteria, taking all comments into consideration.

At the next DUR Board meeting, the Center presents draft final criteria for Board approval. Once approved by the Board, the Center obtains approval from Maryland Medicaid, then it notifies the retrospective and prospective DUR contractors that they may implement the final criteria. The final step in the initial implementation of criteria is for the contractor to inform the Center when the criteria will be (or have been) implemented and any problems regarding criteria which cannot be implemented. Usually the entire approval process takes two to four months. Criteria are then re-evaluated annually based on scientific and clinical advances and an analysis of aggregate DUR screen alert data.

### DUR Criteria for β-agonist Metered Dose Inhalers

The three β-agonist Metered Dose Inhaler (MDI) criteria evaluated in this study were submitted and approved by the DUR Board in May, 1994. First Health then implemented the criteria in July, 1994. (DHMH, 1997).

The first criterion relates to the concurrent use of β-agonist inhalers with non-cardioselective β-blockers. Concurrent use of these two drugs may result in the inhibition of the effects of each drug and is therefore not indicated. Failure to meet this criterion can result in increased severity of symptoms together with increased incidence of emergency treatments and inpatient admissions for respiratory conditions, asthma, emphysema, bronchospasm, bronchitis, COPD, and allied conditions. Exhibit 5.12.1 displays a list of relevant drug products in the two classes.

**Exhibit 5.12.1**

**HCFA DUR Demonstration Evaluation**
**Maryland ProDUR Criteria for Concurrent Use of β-Agonist Inhalers with Non-cardioselective β-blockers**

| β-agonist inhalers | Non-cardioselective β-blockers |
|---|---|
| Albuterol | Carteolol |
| Bitolterol | Nadolol |
| Isoetharine | Penbutolol |
| Isoproterenol | Pindolol |
| Metoproternol | Propranolol |
| Pirbuterol | Sotolol |
| Salmeterol | Timolol |
| Terbutaline | Labetalol |

Source: Maryland DHMH, 1997

The second ProDUR criterion relates to concurrent use of oral and inhaled β-agonists. This combination may result in overdose and excessive beta stimulation and is therefore not indicated. Manifestations of overdose may include anginal pain, hypertension, hypokalemia, palpitations, tachycardia, tremor, nausea, dizziness, heartburn, headache, insomnia, or nervousness. The relevant drug products invoked by this criterion are shown in Exhibit 5.12.2.

**Exhibit 5.12.2**

**HCFA DUR Demonstration Evaluation**
**Maryland ProDUR Criteria for Duplicative Therapy Involving Use of β-agonist Inhalers**

| β-agonist inhalers | Oral β-agonists |
|---|---|
| Albuterol | Albuterol |
| Bitolterol | Isoproterenol |
| Isoetharine | Metoproterenol |
| Isoproterenol | Terbutaline |
| Metoproternol | |
| Pirbuterol | |
| Salmeterol | |
| Terbutaline | |

Source: Maryland DHMH, 1997

Therapeutic duplication involving multiple inhaled β-agonist products was not evaluated. The reason for this is that multiple short-acting products may be prescribed to take advantage of proprietary systems that deliver the drug through the nasal passages or, alternatively, the pharynx. This situation may lead to a high number of false alerts for duplicative therapy.

The third ProDUR criterion evaluated relates to the appropriateness of β-agonist therapy in conjunction with anti-inflammatory agents. The number of inhalers used per month should not exceed that which is listed in Exhibit 5.12.3 below without concurrent use of an oral and/or inhaled anti-inflammatory agent (corticosteroid, cromoglycate, and/or nedocromil). Failure to prescribe anti-inflammatory agents under these circumstances may exacerbate the underlying respiratory condition, thus leading to anginal pain, hypertension, hypokalemia, palpitations, tachycardia, tremor, nausea, dizziness, heartburn, headache, insomnia, or nervousness. The maximum number of inhalers per month listed in Exhibit 5.12.3 was calculated based on the manufacturer's recommended dosing regimen.

The Maryland DUR board considered implementing other β-agonist criteria including drug-disease contraindication and drug-allergy interactions. These criteria were not approved either because the conditions are not identifiable in the claims database or are identifiable only indirectly. It should be noted, however, that the criterion for concurrent use of β-agonist inhalers with non-cardioselective β-blockers does capture individuals with heart disease who are treated with non-cardioselective β-blockers.

The second ProDUR criterion relates to concurrent use of oral and inhaled β-agonists. This combination may result in overdose and excessive beta stimulation and is therefore not indicated. Manifestations of overdose may include anginal pain, hypertension, hypokalemia, palpitations, tachycardia, tremor, nausea, dizziness, heartburn, headache, insomnia, or nervousness. The relevant drug products invoked by this criterion are shown in Exhibit 5.12.2.

**Exhibit 5.12.2**

**HCFA DUR Demonstration Evaluation**
**Maryland ProDUR Criteria for Duplicative Therapy Involving Use of β-agonist Inhalers**

| β-agonist inhalers | Oral β-agonists |
|---|---|
| Albuterol | Albuterol |
| Bitolterol | Isoproterenol |
| Isoetharine | Metoproterenol |
| Isoproterenol | Terbutaline |
| Metoproternol | |
| Pirbuterol | |
| Salmeterol | |
| Terbutaline | |

Source: Maryland DHMH, 1997

Therapeutic duplication involving multiple inhaled β-agonist products was not evaluated. The reason for this is that multiple short-acting products may be prescribed to take advantage of proprietary systems that deliver the drug through the nasal passages or, alternatively, the pharynx. This situation may lead to a high number of false alerts for duplicative therapy.

The third ProDUR criterion evaluated relates to the appropriateness of β-agonist therapy in conjunction with anti-inflammatory agents. The number of inhalers used per month should not exceed that which is listed in Exhibit 5.12.3 below without concurrent use of an oral and/or inhaled anti-inflammatory agent (corticosteroid, cromoglycate, and/or nedocromil). Failure to prescribe anti-inflammatory agents under these circumstances may exacerbate the underlying respiratory condition, thus leading to anginal pain, hypertension, hypokalemia, palpitations, tachycardia, tremor, nausea, dizziness, heartburn, headache, insomnia, or nervousness. The maximum number of inhalers per month listed in Exhibit 5.12.3 was calculated based on the manufacturer's recommended dosing regimen.

The Maryland DUR board considered implementing other β-agonist criteria including drug-disease contraindication and drug-allergy interactions. These criteria were not approved either because the conditions are not identifiable in the claims database or are identifiable only indirectly. It should be noted, however, that the criterion for concurrent use of β-agonist inhalers with non-cardioselective β-blockers does capture individuals with heart disease who are treated with non-cardioselective β-blockers.

**Exhibit 5.12.3**

**HCFA DUR Demonstration Evaluation**
**Maryland ProDUR Criteria for Therapeutic Use of β-agonist Inhalers**

| β-agonist inhalers | Maximum number of inhalers per month |
|---|---|
| Albuterol | 2 |
| Bitolterol | 2 |
| Isoetharine | 2 |
| Isoproterenol | 2 |
| Metoproternol | 2 |
| Pirbuterol | 2 |
| Pirbuterol (Autoihaler) | 1 |
| Salmeterol | 1 |
| Terbutaline | 2 |

Source: Maryland DHMH, 1997

## Adverse Clinical Outcomes Related to β-agonist Therapy

The pharmacological actions of β-agonist MDIs have been well investigated and established to be effective in the treatment of asthma and other respiratory conditions, such as COPD and emphysema. β-agonist MDIs cause a stimulation of β-2 receptors, which are primarily located in the respiratory tract, thereby providing a dose-dependent relaxation of the bronchial smooth muscle. Unfortunately, currently available β-agonists stimulate β-1 as well as β-2 receptors. Both β-1 and β-2 receptors are located in the myocardium. Regardless of the original medical need for β-agonist MDI use, the mechanism of action and potential side effects are the same for all patients. Sensitivity to side effects is partially dose-related but it is also influenced by the patients' health status and inhalation techniques. The same range of side effects is possible for all patients irrespective of age [Beasley (1993); Clauzel (1994); Spangler (1989)].

Because the pharmacological action of β-agonists are dose-dependent, many of the side effects that occur are partially related to an over stimulation of the β-receptors. β-1 receptors occur in a variety of different organ systems so there is a range of potential side effects. It has been found that tachycardia and palpitations are the most common side effects of β-agonist inhalers. These side effects occur by causing a vasodilation of the peripheral vasculature and stimulation of the β-receptors on the heart. The peripheral vasodilation can also cause dizziness and hypertension. Skeletal muscle β-receptor stimulation is a common occurrence with β-agonist MDIs and over stimulation can result in tremors [Beasley (1993); Clauzel (1994); Spangler (1989)].

Hypokalemia may occur because bronchodilation and stimulation of cells requires energy, which requires the transportation of potassium across cellular membranes. This in turn can reduce blood potassium levels. With chronic use of β-agonists most asthmatics will develop a tolerance for hypokalemia. Those asthmatics at greatest risk for adverse outcomes have certain pre-existing conditions, such as ischemic heart disease or hypokalemia, or are taking drugs such as oral corticosteroids, theophyllines, diuretics or digoxin. Hypokalemia in turn can produce arrhythmias and anginal pain in effected asthmatics [Beasley (1993); Clauzel (1994); Spangler (1989)].

β-agonist agents are also known to cause stimulation of the central nervous system producing insomnia, nervousness and headaches. In most cases these side effects are troublesome but can be dealt with by adjusting the medication or the dosing schedule. Caution is warranted for individuals who suffer from migraine or vascular headaches. There has been limited study about how these side effects occur [Beasley (1993); Clauzel (1994); Spangler (1989)].

As part of the clinical evaluation of all patients with asthma and other chronic respiratory conditions, physicians must determine the most effective dose of β-agonists with the fewest side effects. In those cases where respiratory control cannot be established, the introduction of an inhaled corticosteroid can decrease the overall dose of the β-agonist and achieve a reduction in acute exacerbations. By decreasing the overall dose there is a decreased risk of β-receptor over-stimulation and, therefore, a possible reduction in β-agonist induced side effects. Corticosteroids work through an anti-inflammatory effect that reduces airway responsiveness to both indirect and direct stimuli [Barnes (1993); Clauzel (1994); Fahey (1995)].

Non-cardioselective β-blocking drugs, such as propranolol and pindolol, are used in the treatment of many different conditions including angina, hypertension, arrhythmias, and migraines after myocardial infarctions. For patients suffering from chronic respiratory conditions, these drugs may precipitate or worsen bronchoconstriction due to the concurrent stimulation of both β-1 and β-2 receptors in the heart and bronchial passages. The more recently developed β-blockers have extensive cardioselectivity, so they are less likely to induce bronchospasm [Giulekas (1986); Patakas (1983); Prichard (1986); Rosenow (1994)].

### Methods

This evaluation was designed to determine if Maryland's ProDUR screening program for β-agonist MDIs had an impact on the incidence of selected adverse clinical outcomes related to misuse of these medications. It was hypothesized that alerts for potential misuse of β-agonist inhalers would prompt pharmacists to intervene to correct the problems and thereby reduce the incidence of the adverse consequences in the period following implementation of the screening program. To test this hypothesis, the study compared incidence rates for adverse outcomes for two-and-a-half years following the implementation of the β-agonist screens (the "follow-up" period) in Maryland, Iowa, and Georgia using regression analysis to control for population differences in the year prior to the ProDUR intervention (the "baseline" period). Maryland was the only state of the three to have implemented on-line DUR screening for of β-agonist inhalers during the study period. Georgia has not implemented OPDUR in any form. Iowa's OPDUR program operated as a HCFA demonstration site during this period, and did not target β-agonist inhalers

***Study Population.*** The study population consisted of elderly dual-eligible Medicaid recipients in the three states who met the following criteria at baseline:

- 65 or older as of July 1993 with no nursing home claims or indicator of HMO enrollment;
- Evidence of treatment for one of the following conditions: asthma (ICD-9 codes 493-508.9), emphysema (510-510.9), bronchospasm (519.1), bronchitis (466-466.1), COPD and allied conditions (490-492), or other respiratory conditions (786-786.09);

- Also included were persons with no pulmonary ICD-9 who had one or more claims for any of the following drugs: cromones (disodium cromoglycate; nedocromil), inhaled steroids (dexamethasone sodium phosphate, flunisolide, triamcinolone acetonide, beclomethasone dipropionate, fluticasone propionate), beta-agonists (metaproteranol, bitolterol, isotharine, isoproteranol, terbutaline sulfate, albuterol, salmeterol), methylxanthine (theophylline, oxtriphylline), anticholinergic (ipratropium bromide – Atrovent).

***Data and Model Variables.*** As in the other outcomes studies, data were extracted from Medicaid drug claims and Medicare physician and hospital bills for the 42 months of the study. The data elements included measures for the dependent variables shown in Exhibit 5.12.4. These dependent variables were selected by an expert clinical based on the literature summarized in the preceding section. Other model variables were selected for the following domains: intervention exposure, basic demographic characteristics, and risk factors specific to each hypothesized outcome.

**Exhibit 5.12.4**

**HCFA DUR Demonstration Evaluation**
**Adverse Clinical Outcomes Specified as Dependent Variables**

| Variable | ICD-9 Classification |
|---|---|
| Asthma | 493-508.9 |
| Emphysema | 510-510.9 |
| Bronchospasm | 519.1 |
| Bronchitis | 466-466.1 |
| COPD and allied conditions | 490-492 |
| Other respiratory | 786-786.09 |
| Anginal pain | 413.9 |
| Hypertension and elevated blood pressure | 401-405.9, 796.2 |
| Palpitations/tacycardia | 785.1, 427-427.2 |
| Mixed symptom cluster including: tremor, nausea, dizziness, heartburn, headache insomnia or nervousness | 781.0, 787.02, 780.4, 787.1, 784.0, 780.52, 799.2 |

Source: International Classification of Diseases, 9tn Revision, DHHS Publication No. (PHS) 91-1260, October, 1991.

***Intervention Exposure.*** As noted above, Maryland was the only state of the three to have targeted β-agonist inhalers for prospective DUR screening during the study period. Hence exposure to the intervention is defined as being a Maryland Medicaid recipient during the follow-up period. The follow-up period for each study subject began when the individual filled his/her first prescription after August 1, 1994 (the month following the intervention month in Maryland). If the subject lost eligibility for Medicare or Medicaid, entered a nursing home, entered an HMO, or died, the follow-up period was concluded at that point. Because the probability that subjects incurred both disease-based risk factors and outcomes are positively associated with the length of the baseline and follow-up periods, the estimating models include these measures.

*Demographic Variables.* The Medicaid enrollment and claims files contain very limited data on recipients' personal and demographic characteristics. Three sets of variables were used: age categorized by five-year intervals (65-69, 70-74, 75-79, 80-84, 85+), race categorized as white, black or other, and gender.

*Risk Factors.* Two types of risk factors were identified as relevant control variables for the regression analyses: 1) comorbid conditions that might either influence the form of β-agonist therapy or the outcomes of using β-agonists, and 2) drugs (other than β-agonists) that might also covary with therapeutic decisions or clinical outcomes of interest to the study. A list of the risk factors is provided in Exhibit 5.12.5.

**Exhibit 5.12.5**

**HCFA DUR Demonstration Evaluation**
**Relevant Risk Factors for the Study of β-agonist Outcomes**

| Variable | ICD-9 Classification |
|---|---|
| *Clinical conditions* | |
| Hypertensive cardiac w/CHF | 402.91 |
| Ischemic heart disease | 410-414 |
| Atherosclerosis | 440.9 |
| Cardiac disrhythmias | 427 |
| Heart Failure | 428 |
| Hypokalemia | 276.8 |
| Diabetes | 250 |
| *Drugs* | |
| Corticosteroids | |
| Antihistamines | |
| Theophyline | |
| Oral beta agonists | |
| Inhaled beta agonists | |
| Non-cardioselective blockers | |
| NSAIDs | |
| Cromolyn sodium | |
| Nedocromil | |

Source: International Classification of Diseases, 9tn Revision, DHHS Publication No. (PHS) 91-1260.

### Model Specification

The statistical methods employed to test the study hypotheses are similar to those used in the other outcome studies reported in Sections 4.8-4.11 and 5.8-5.11. The basic regression specification is the same: the dependent variables are binary indicators of whether the subject had evidence of the outcome in question during the follow-up period; a ProDUR treatment effect is captured by a dummy variable (Maryland Medicaid recipient =1; 0 otherwise); baseline risk factors and demographic characteristics are represented by binary variables; the length of the baseline and follow-up periods are entered in log form as is the maximum number of drug GPIs utilized in the baseline period. Because this is a cross-state rather than a within-state study, there were no concerns related to pharmacy clusters or contamination by patients visiting treatment and control pharmacies. However, the cross-state design raised another important concern. Since subjects were not randomized, how can one tell whether a finding or a treatment effect is really associated with the intervention itself or some other unobserved factor that covaries with the treatment exposure variable and the outcome measures? The only choice the modeler has under these circumstances is to carefully specify covariates that should, in theory, capture the impact of these unobserved influences. The risk factors are designed to do just that, but to strengthen the specification further, baseline values for all of the outcomes variables were entered as covariates in one set of models. In all, twenty regressions were estimated, ten each for Maryland compared to Georgia and ten for Maryland compared to Iowa.

### Findings

Descriptive findings are summarized in Exhibits 5.12.6 through 5.12.19. The characteristics of the study subjects at baseline are shown in Exhibit 5.12.6 and 5.12.7. As can be seen, the relative sizes of the state samples vary considerably from 4,133 study subjects in Iowa to 24,061 in Georgia. The samples were quite similar in age and gender, but differed radically in racial composition. Only 6 percent of the elderly Iowa subjects were nonwhite compared to 53 percent in Georgia and 48 percent in Maryland. There are also important differences in sample mortality (13 percent of subjects died in the follow-up period in Maryland compared to 19 percent in Iowa and 24 percent in Georgia). Despite these differences, the average length of the follow-up periods were quite similar, ranging from 20 to 24 months in duration.

**Exhibit 5.12.6**

**HCFA DUR Demonstration Evaluation**
**Characteristics of the Three Study Samples of**
**Dual-Eligible 65+ Medicaid Recipients in Maryland, Georgia, and Iowa**

| | Georgia | | Iowa | | Maryland | |
|---|---|---|---|---|---|---|
| **Sample Sizes** | 24,061 | | 4,133 | | 10,075 | |
| *Demographic characteristics* | Mean | SD | Mean | SD | Mean | SD |
| Age as of July 1993 | 75.5 | 7.2 | 75.4 | 6.9 | 74.8 | 6.9 |
| Gender: % male | 24.0 | 0.4 | 22.6 | 0.4 | 20.1 | 0.4 |
| Race: % white | 47.2 | 0.5 | 94.0 | 0.2 | 51.9 | 0.5 |
| Race: % black | 36.4 | 0.5 | 2.7 | 0.2 | 39.9 | 0.5 |
| Race: % other race | 16.4 | 0.4 | 3.3 | 0.2 | 8.2 | 0.3 |
| *Observation period* | | | | | | |
| Baseline months | 11.5 | 1.9 | 10.9 | 2.6 | 11.2 | 2.2 |
| Follow-up months | 23.6 | 9.4 | 19.9 | 10.9 | 22.7 | 9.7 |
| Mortality rate: % dying during observation period | 25.2 | 0.6 | 19.4 | 0.4 | 13.4 | 0.3 |

Source: Medicaid drug data from Maryland, Georgia, and Iowa, 1993-1996

The data in Exhibit 5.12.7 suggest that the three samples of elderly Medicaid recipients had roughly similar clinical risk profiles. Between 20 and 35 percent of each sample had evidence in baseline of ischemic heart disease, cardiac disrhythmias, heart failure, and diabetes. Hypertensive cardiac conditions w/CHF, atherosclerosis and hypokalemia are all relatively rare with baseline incidence rates below ten percent in each sample. While the differences are small, the Iowa sample had the lowest incidence of comorbid conditions of the three. On the other hand, the highest reported utilization of drug risk factors were in Iowa (see the bottom panel of Exhibit 5.12.7). The lowest rate of drug risk factors occurred in Georgia.

**Exhibit 5.12.7**

### HCFA DUR Demonstration Evaluation
### Incidence of Treatment for Comorbid Conditions and Use of Selected Drugs at Baseline by Dual-Eligible 65+ Medicaid Recipients in Maryland, Georgia, and Iowa

| | Baseline Incidence Rates and Counts of Prescriptions | | | | | |
|---|---|---|---|---|---|---|
| | Georgia | | Iowa | | Maryland | |
| *Clinical conditions (%)* | Mean | SD | Mean | SD | Mean | SD |
| Hypertensive cardiac w/CHF | 2.12% | 0.14 | 0.65% | 0.08 | 2.81 | 0.17 |
| Ischemic heart disease | 33.22 | 0.47 | 28.38 | 0.45 | 35.24 | 0.48 |
| Atherosclerosis | 2.91 | 0.17 | 2.34 | 0.15 | 2.90 | 0.17 |
| Cardiac disrhythmias | 20.56 | 0.41 | 17.78 | 0.38 | 24.90 | 0.43 |
| Heart Failure | 27.84 | 0.45 | 25.36 | 0.44 | 28.00 | 0.45 |
| Hypokalemia | 9.3 | 0.29 | 8.06 | 0.27 | 7.35 | 0.26 |
| Diabetes | 27.10 | 0.44 | 22.11 | 0.42 | 26.81 | 0.44 |
| *Drug use (number of fills)* | | | | | | |
| Corticosteroids | 0.72 | 2.10 | 0.57 | 2.18 | 0.84 | 2.56 |
| Antihistamines | 1.02 | 2.47 | 1.05 | 2.92 | 0.48 | 1.56 |
| Theophyline | 0.89 | 2.54 | 2.47 | 6.82 | 0.91 | 2.58 |
| Oral beta agonists | 0.37 | 1.74 | 0.92 | 3.68 | 0.45 | 2.18 |
| Inhaled beta agonists | 1.08 | 3.22 | 2.21 | 5.54 | 1.30 | 3.73 |
| Non-cardioselective blockers | 0.55 | 2.07 | 0.60 | 2.17 | 0.52 | 1.85 |
| NSAIDs | 1.52 | 2.77 | 3.74 | 7.55 | 2.23 | 5.01 |
| Cromolyn sodium | 0.01 | 0.26 | 0.11 | 1.37 | 0.04 | 0.67 |
| Nedocromil | 0.00 | 0.12 | 0.01 | 2.92 | 0.00 | 0.15 |

Source: Medicaid drug data from Maryland, Georgia, and Iowa, 1993-1996

Exhibit 5.12.8 presents baseline and follow-up period incidence rates for the ten adverse clinical outcomes selected for the study. In every case the incidence rates are higher in the follow-up period. This should not be taken to mean that the populations are necessarily sicker, but rather that the follow-up observation periods are longer and therefore the probability of incurring an adverse event is higher. There is also considerable interstate variability evident in these data. In most cases, the incidence rates are highest in Maryland and lowest in Iowa.

**Exhibit 5.12.8**

### HCFA DUR Demonstration Evaluation Incidence of Treatment for Selected Clinical Outcomes Associated With Inappropriate Use of β-agonist Inhalers By Dual-Eligible 65+ Medicaid Recipients in Maryland, Georgia, and Iowa

| | Maryland | | Georgia | | Iowa | |
|---|---|---|---|---|---|---|
| | **Baseline** | **Follow-up** | **Baseline** | **Follow-up** | **Baseline** | **Follow-up** |
| *Mean observation period in months* | 11.2 | 22.7 | 11.5 | 23.6 | 10.9 | 19.9 |
| Clinical outcome | | | | | | |
| Asthma | 41.11% | 47.98% | 40.98% | 45.05% | 40.50% | 47.30% |
| Emphysema | 0.22 | 0.53 | 0.17 | 0.25 | 0.02 | 0.07 |
| Bronchospasm | 1.51 | 2.94 | 1.50 | 2.53 | 1.11 | 2.15 |
| Bronchitis | 22.40 | 26.25 | 27.32 | 28.40 | 29.06 | 28.33 |
| COPD | 28.86 | 35.28 | 29.50 | 33.41 | 22.33 | 25.77 |
| Other respiratory | 35.28 | 44.79 | 30.53 | 39.87 | 19.99 | 30.83 |
| Anginal pain | 12.93 | 19.50 | 11.06 | 17.21 | 8.83 | 34.51 |
| Hypertension | 64.52 | 74.54 | 63.23 | 74.23 | 39.46 | 49.81 |
| | 6.57 | 11.67 | 5.02 | 9.53 | 3.10 | 6.00 |
| Palpitations/tacycardia | | | | | | |
| Mixed symptom cluster | 13.41 | 25.37 | 18.15 | 31.49 | 10.21 | 41.31 |

Source: Medicaid drug data from Maryland, Georgia, and Iowa, 1993-1996

The ten charts numbered 5.12.9 through 5.12.18 depict monthly incidence rates for each adverse clinical outcome beginning with August 1993 and ending with December 1996. The implementation of Maryland's ProDUR screening for β-agonist inhalers is shown by the vertical line above July 1994. Exhibit 5.12.9 shows that monthly incidence of treatment for asthma among members of the study samples hovered around 11 percent across the states with no evident pattern in the follow-up period. Treatment for emphysema is much less common (Exhibit 5.12.10) with modal values of zero for the Iowa sample and under 8/100 of a percent in the other two states. Here too there are neither evident time trends nor suggestions of an intervention effect in Maryland. The incidence of bronchitis (Exhibit 5.12.11) is highly seasonal in each state with peaks during the winter months. The worst problems were evident in the winter of 1993. Again, there is no obvious treatment effect for the Maryland sample. The incidence of bronchospasm is rare in all states (Exhibit 5.12.12), albeit Maryland rates are slightly higher than Georgia or Iowa overall. The incidence of COPD and allied conditions (Exhibit 5.12.13) is significantly higher in the Maryland sample over the entire study period with rates double those in Iowa. No ProDUR impact is evident here. Exhibit 5.12.14 presents trends in the incidence of all other respiratory conditions combined. Here too, Maryland rates are higher overall, but there is a drop after a peak of six percent in February and March 1994. The pattern of high incidence rates in Maryland is repeated in Exhibit 5.12.15 for anginal pain. In the baseline period, Maryland's rate was double that of Iowa (20-25 percent versus 10-16 percent). The incidence of anginal pain converge somewhat during the follow-up period, but Maryland remains the highest of the three states. Exhibits 5.12.16 and 5.12.17 show trends in the incidence of treatment for hypertension/elevated blood pressure and palpitations/tachycardia, respectively. Once again, Maryland's rates are higher than the other two states throughout the study period for both outcomes, and there are no evident intervention effects. The final chart (Exhibit 5.12.18) shows monthly incidence rates for the mixed symptom cluster (see definition in Exhibit 5.12.4 above). Here Maryland's rates fall between Georgia and Iowa, but they are higher port-intervention—not lower as hypothesized.

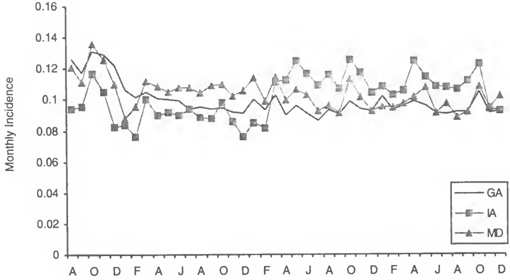
Although there is no obvious pattern of reduced incidence of adverse outcomes in Maryland after July 1994 relative to Iowa and Georgia, these descriptive profiles do not control for differences in demographic characteristics (Exhibit 5.12.6) and risk factors (Exhibit 5.12.7) among the three samples of elderly Medicaid recipients. Conditional estimates of the impact of Maryland's β-agonist ProDUR program on elderly Medicaid recipients are shown in Exhibit 5.12.19. This table presents marginal probabilities, probit coefficients and standard errors for the intervention effect dummy variables from each of the 20 probit regressions.[54] The marginal probabilities can read as percentage differences in the probability of the adverse outcome in the follow-up period between each pair of states. Thus, the marginal probability of 0.035 in the asthma regression for Maryland versus Georgia means that Maryland subjects are 3.5 percent more likely to have the adverse outcome than their Georgia counterparts, all else being equal. It will be noted that all but three treatment effect coefficients are positive (not negative as hypothesized), and most are highly significant even when the predicted effect is very small.[55]

---

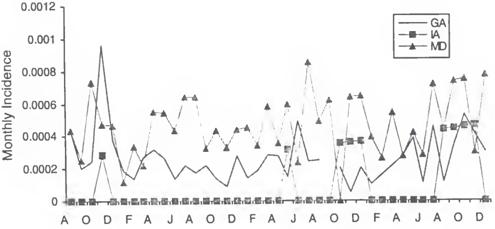
[54] The coefficient shown are from the 20 models estimated with baseline values of the outcomes variables entered as covariates. Another set of regressions was estimated excluding these variables. The results were similar in signs and significance levels and hence are not reported.

[55] Because of the large sample sizes used for the analysis, even small effects are likely to be statistically significant.

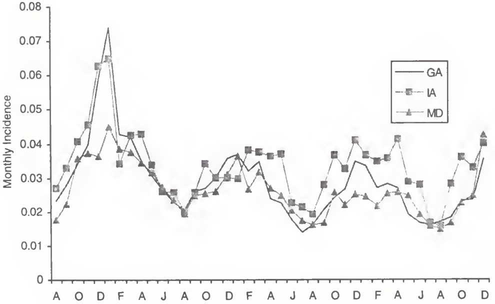
## Exhibit 5.12.9
## HCFA DUR Demonstration Evaluation
## Monthly Incidence Rates for Treatment of Asthma Among Members of the Beta Agonist Study Samples in Maryland, Georgia and Iowa, 1993-1996



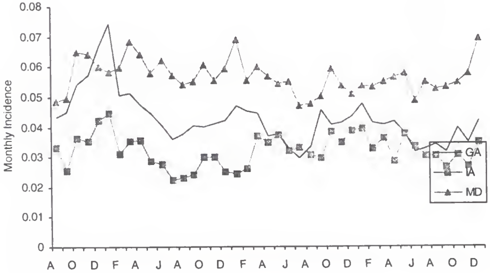
## Exhibit 5.12.10
## HCFA DUR Demonstration Evaluation
## Monthly Incidence Rates for Treatment of Emphysema Among Members of the Beta Agonist Study Samples in Maryland, Georgia and Iowa, 1993-1996



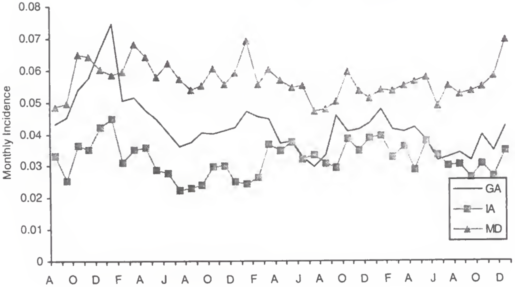
## Exhibit 5.12.11
## HCFA DUR Demonstration Evaluation
## Monthly Incidence Rates for Treatment of Bronchitis Among Members of the Beta Agonist Study Samples in Maryland, Georgia and Iowa, 1993-1996



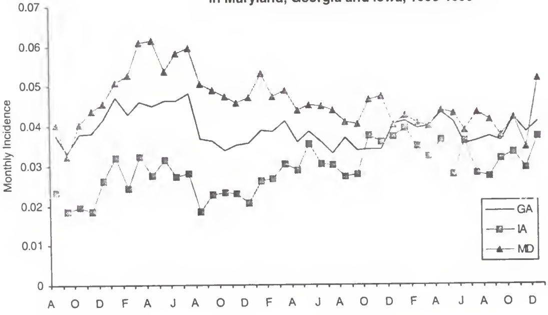
## Exhibit 5.12.12
## HCFA DUR Demonstration Evaluation
## Monthly Incidence Rates for Treatment of Bronchospasm Among Members of the Beta Agonist Study Samples in Maryland, Georgia and Iowa, 1993-1996



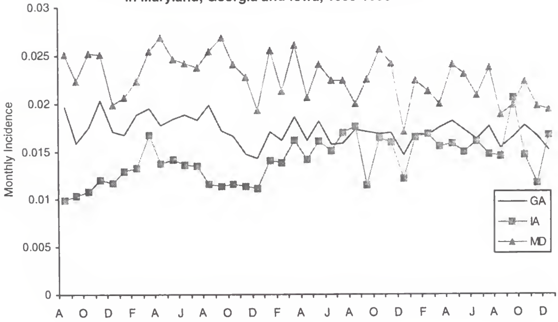
### Exhibit 5.12.13
### HCFA DUR Demonstration Evaluation
### Monthly Incidence Rates for Treatment of COPD and Allied Conditions Among Members of the Beta Agonist Study Samples in Maryland, Georgia and Iowa, 1993-1996



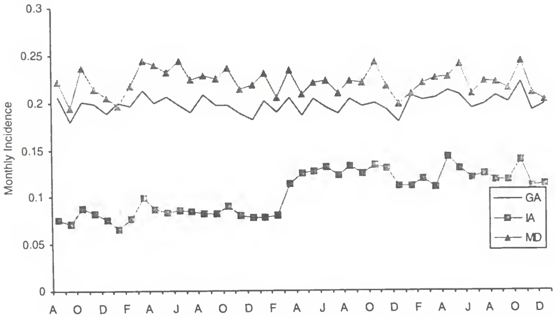
### Exhibit 5.12.14
### HCFA DUR Demonstration Evaluation
### Monthly Incidence Rates for Treatment of Other Respiratory Conditions Among Members of the Beta Agonist Study Samples in Maryland, Georgia and Iowa, 1993-1996



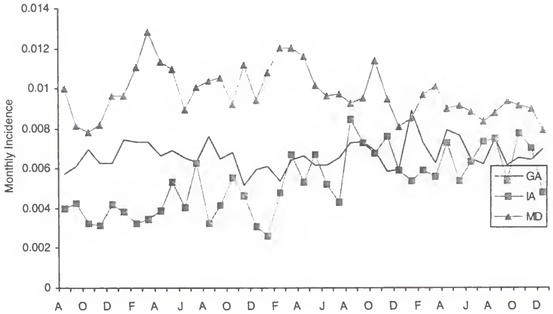
## Exhibit 5.12.15
## HCFA DUR Demonstration Evaluation
## Monthly Incidence Rates for Treatment of Anginal Pain Among Members of the Beta Agonist Study Samples in Maryland, Georgia and Iowa, 1993-1996



## Exhibit 5.12.16
## HCFA DUR Demonstration Evaluation
## Monthly Incidence Rates for Treatment of Hypertension and Elevated BP Among Members of the Beta Agonist Study Samples in Maryland, Georgia and Iowa, 1993-1996



## Exhibit 5.12.17
## HCFA DUR Demonstration Evaluation
## Monthly Incidence Rates for Treatment of Palpitations and Tachycardia Among Members of the Beta Agonist Study Samples in Maryland, Georgia and Iowa, 1993-1996



## Exhibit 5.12.18
## HCFA DUR Demonstration Evaluation
## Monthly Incidence Rates for Treatment of Mixed Symptom Among Members of the Beta Agonist Study Samples in Maryland, Georgia and Iowa, 1993-1996

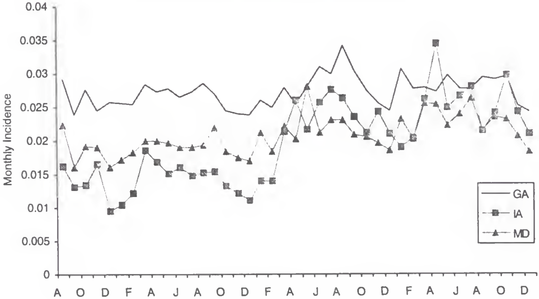


Exhibit 5.12.18

## HCFA DUR Demonstration Evaluation
## Regression Results of the Impact of the Maryland OPDUR Intervention for Beta Agonist Inhaler Use by Dual-Eligible 65+ Medicaid Recipients

| | Maryland versus Georgia | | | Maryland versus Iowa | | |
|---|---|---|---|---|---|---|
| | Marginal probability | Probit coefficient | Standard error | Marginal probability | Probit coefficient | Standard error |
| Asthma | 0.035** | 0.087 | 0.018 | 0.085** | 0.219 | 0.033 |
| Emphysema | 0.002** | 0.247 | 0.075 | 0.002* | 0.439 | 0.214 |
| Bronchospasm | 0.004* | 0.073 | 0.035 | 0.011** | 0.168 | 0.064 |
| Bronchitis | 0.005 | 0.016 | 0.018 | 0.007 | 0.020 | 0.033 |
| COPD | 0.041** | 0.109 | 0.018 | 0.152** | 0.395 | 0.033 |
| Other respiratory | 0.061** | 0.154 | 0.017 | 0.152** | 0.384 | 0.031 |
| Anginal pain | 0.030** | 0.120 | 0.021 | 0.054** | 0.208 | 0.039 |
| Hypertension | 0.019** | 0.073 | 0.021 | 0.083** | 0.301 | 0.034 |
| Palpitations/tacycardia | 0.022** | 0.122 | 0.023 | 0.055** | 0.289 | 0.045 |
| Mixed symptom cluster | -0.035 | -0.105 | 0.018 | 0.037** | 0.117 | 0.034 |

Source: Medicaid drug data from Maryland, Georgia, and Iowa, 1993-1996.
Notes: Statistical significance indicated by * if p<.05 and ** if p<.01 in two-tailed tests.

There is no ready explanation for these counterintuitive findings. Clearly, the analysis provides no support for the study hypotheses that ProDUR screening led to improved clinical outcomes for elderly Maryland Medicaid recipients with chronic respiratory problems. The positive signs on the treatment effect variables suggest that some other unmeasured risk factor negatively influenced the pattern of outcomes in Maryland compared to the other two states. What that negative influence might be is unknown. There were no overt policy changes in the Maryland Medicaid program during this period that could explain it.

One might argue that the regression findings are statistical artifacts due either to the time period selected for the study or the choice of statistical model. It is true that the incidence rates for several of the outcome measures were rising (e.g., bronchospasm, palpitations and tachycardia, other respiratory conditions) or falling (asthma) during the baseline period. Because the measured effect of the intervention is sensitive to history, changing the length of the baseline time period could conceivably alter the findings. Likewise, for some outcomes (e.g. anginal pain, palpitations and tachycardia, other respiratory conditions) the incidence rates appear to converge in the follow-up period, with Maryland rates falling relative to the other states. A cross-sectional pre-post design like that employed in this study is less sensitive to trends than an interrupted time-series design. For these conditions it is possible that an ARIMA model might produce treatment effects in the hypothesized direction. However, it must be emphasized that these are selective ex-post observations. It is extremely unlikely that either analytical procedure—lengthening the baseline or using ARIMA modeling—would change the study findings across the entire set of outcome variables.

## 5.13 Costs and Cost-Effectiveness of the OPDUR Demonstration

### Measuring the Costs and Benefits of the Iowa OPDUR Program

Thus far, this study has assessed the effectiveness of the Iowa OPDUR system. It was hypothesized that the system would help pharmacists dispense more appropriately (leading to fewer drug-related problems), reduce medical services utilization, and reduce costly prescriptions. Previous sections have addressed the impact of the Iowa OPDUR system on prescription drug utilization, downstream Medicare and Medicaid utilization, and a variety of clinical outcomes related to specific drug interactions.

In this section, the analysis centers on assessing the cost of OPDUR. Funds allocated to the Iowa OPDUR program could have been allocated to other projects. The *opportunity cost* of the OPDUR program is the forgone benefits that could have been achieved had this alternative project been implemented. To allocate scarce resources efficiently, it is useful to explicitly measure and value the costs and benefits of alternative projects. This process is called *economic evaluation* [Weisbrod (1994)]. Because the analyses presented in the preceding chapters did not produce definitive evidence that OPDUR was effective, this chapter focuses on the costs of the intervention and makes no attempt to combine this information with data on benefits.

The objective of this analysis is to use Iowa OPDUR demonstration budgets to estimate: (1) the cost of operating an "enhanced" OPDUR program (similar to the Iowa demonstration's OPDUR program); and (2) the cost of operating a "basic" OPDUR program (similar to those that currently exist in other states). A distinction was made between these two costs because Iowa OPDUR was a state-of-the-art program with many enhancements that usually are not included in the more basic OPDUR programs implemented in other states. The average operating cost of OPDUR per year and the average operating cost of OPDUR per prescription also are computed for the basic and the enhanced models. These ratios might be useful to another state implementing a similar program.[56]

### Data

The data used in this analysis come from Iowa OPDUR demonstration budgets for the time period 1992-1997. The budgets were provided by the University of Iowa. Each budget is divided into three sections: Iowa, Iowa Pharmacists Association and the University of Iowa. The sections contain line items that were budgeted for each of these entities.

The Heath Care Financing Administration signed an agreement for the OPDUR demonstration with the Iowa Department of Human Resources, in cooperation with researchers from the University of Iowa, Drake University, and the Iowa Pharmacists Association. In turn, Iowa selected the National

---

[56] It also would be interesting to compare OPDUR costs across various programs in different states. Differences in OPDUR costs across states may reflect the varying efficiency of OPDUR programs, differences in the quality of the OPDUR system, or economies of scale/scope. For example, states with well-developed claims processing systems may find it easier and less costly to implement OPDUR relative to states with less developed claims processing systems. Nevertheless, because available data is limited to the Iowa OPDUR program. this analysis cannot address other states' OPDUR program costs.

Data Corporation (NDC) to serve as the vendor of the OPDUR software. The Medicaid fiscal intermediary, Unisys, served as the host of the OPDUR software.

The University of Iowa also provided detailed information on actual demonstration expenditures during this time period. This expenditure information was very useful, because in some cases less or more than the budgeted amount was spent. Unfortunately, expenditure information was not available for the state of Iowa and the Iowa Pharmacists Association portions of the budget.

## Methodology

The estimation of OPDUR costs involved many assumptions and methodological decisions. The sections below outline the viewpoint taken in the study, assumptions made about research and operations costs, and treatment of capital, indirect and start-up costs.

### *Viewpoint of the Analysis*

The types of costs included in a study may vary by the perspective that explicitly or implicitly underlies the analysis. For example, if pharmacists spend large amounts of time learning how to use the OPDUR system, this time is a cost from the perspective of the pharmacist, but it may not be a cost from the perspective of the government, the prescriber or the patient. In this analysis, a government payor viewpoint was adopted; costs were included only if they were borne by government.

Ideally, the societal viewpoint would be most appropriate for this study because it is society as a whole that bears the costs and receives the benefits of the OPDUR system; the societal viewpoint implies that all costs, regardless of who incurs them, should be included in the analysis.[57] The societal viewpoint, however, suggests that a variety of cost data should be collected, including information from demonstration staff, patients, prescribers and pharmacists, and data on time costs and costs of donated materials. This analysis was limited to information available in the demonstration budgets. It was expected that other costs would be fairly incidental.[58] Nevertheless, because no information was available about these types of costs, the societal perspective could not be adopted in this analysis.

### *Research and Operations Costs*

The demonstration budgets combine funds used for research purposes and funds used for OPDUR operations. One of the central challenges of this analysis was to separate research costs from operations costs. This task was especially difficult for staff line items, since many of the demonstration staff members participated in both research activities and OPDUR operations as needed.

---

[57] It is true, however, that some groups bear more costs than others, and some groups enjoy OPDUR benefits more directly than others. For example, a taxpayer who is not dually eligible for Medicare and Medicaid incurs OPDUR costs, but receives benefits only indirectly, through the improved health outcomes of others.

[58] For example, the OPDUR program did not involve large quantities of donated goods or time, and it is unlikely that patients incurred any significant costs related to the demonstration. Pharmacists, and to a much lesser extent, prescribers, did spend time learning about the OPDUR program, but it is unlikely that this training took considerable amounts of time away from other work or leisure time.

Dr. Elizabeth Chrischilles, principal investigator at the University of Iowa, estimated the proportion of each line item that was used for operational activities as opposed to research. For some budgeted line items, such as cognitive services reimbursement and pharmacy recruitment, it was clear that no portion of the funds was used for OPDUR operations. In these cases, the entire line item was excluded from the analysis. Otherwise, the proportions used for operations varied from year to year and ranged from .1 to 1. (See Appendix J, proportion operations column)

The estimated proportions reflect the principal investigator's extensive knowledge of all aspects of the demonstration. Nevertheless, there is no practical way to actually measure the proportion of each line item that was used for OPDUR operations. Ideally, one would like to obtain estimates from a variety of key staff members, and then test the sensitivity of the results to changes in the estimated proportions. This effort was beyond the scope of this analysis. Although a general sensitivity analysis was performed, the sensitivity of the results to changes in the proportion estimates was not tested.

### *Budgeted Versus Actual Expenditures*

The demonstration budget line items reflect funds that were granted by the Health Care Financing Administration. The funds may have been re-allocated between line items or not spent at all. This event would affect the analysis, which is based on extracting the OPDUR operations component from each line item. In particular, if certain line items were not spent, the cost of operating the OPDUR program might be over-estimated.

The University of Iowa provided actual expenditures for each line item for every year of the demonstration. The University of Iowa line items constitute the majority of the budgets in every year of the demonstration. Expenditure information was not available for the state of Iowa and the Iowa Pharmacists Association portions of the budget. For the state of Iowa and the Iowa Pharmacists Association, therefore, it is assumed that what was budgeted is a close approximation of what was actually spent.[59]

### *Annuitization of Capital Costs*

Implementation of OPDUR required initial purchases of capital, including computers and office equipment. For these durable items, the purchase cost was annuitized over the estimated life of the asset. Annuitization generates a yearly rental rate that incorporates both depreciation and the opportunity cost of funds invested in the capital [Weisbrod (1994)]. It was assumed that the computers lasted for five years, the length of the demonstration. Office equipment was assumed to last ten years. A five percent discount rate was used in these calculations.[60]

---

[59] Ms. Marsha Stark at Iowa Medicaid confirmed that this assumption was reasonable for the state's portion of the budget. Two budgeted state expenditures for support pharmacist and for manuals for pharmacists were excluded from the analysis because one of the principal investigators, Dr. Elizabeth Chrischilles, noted that these line items were never provided.

[60] The use of a five percent discount rate was conservative. Often, a discount rate of three percent is used when evaluating public sector projects .

### *Enhanced and Basic OPDUR Systems*

The Iowa OPDUR system had many enhanced, state-of-the-art features that the average state OPDUR program would not include. Consequently, it was necessary to make two OPDUR calculations: (1) the cost of an "enhanced" Iowa OPDUR system; and (2) the cost of a more "basic" OPDUR system, similar to those currently in operation in other states. Compared to the basic OPDUR system, the enhanced OPDUR system costs included a higher proportion of programmer costs at the state level, a corresponding higher level of indirect costs (as described below), more long-distance phone calls, and more staff time during the start-up period.[61]

### *Indirect Costs*

The Iowa OPDUR demonstration was primarily based at the University of Iowa. The indirect costs of the demonstration therefore were calculated using rates set by the university. University indirect cost rates tend to be quite high, and they are not likely to be a good approximation of the indirect costs that would be incurred by a state implementing an OPDUR program. Furthermore, many of the indirect costs incurred in the OPDUR demonstration may have been related to research rather than operations.

To generate an estimate of the indirect costs associated with the *operations portion* of the Iowa OPDUR program, it was assumed that indirect costs incurred are directly proportional to direct costs incurred. The following formula was used to calculate indirect costs for the basic OPDUR program:

[(basic direct costs attributable to operations)/(all direct costs)]x indirect costs

A similar formula was used to calculate indirect costs for the enhanced OPDUR program:

[(enhanced direct costs attributable to operations)/(all direct costs)]x indirect costs

Because the direct costs attributable to operations is somewhat larger for the enhanced OPDUR system compared to the basic OPDUR system, the indirect costs are also somewhat higher for the enhanced system compared to the basic system.

### *Issues Related to Start-Up Costs and Time Frame*

The Iowa OPDUR demonstration project took place over five years, September 30, 1992 - September 29, 1997. The OPDUR system was actually in operation, however, from June 1994 - June 1997. The time before and after the operations period was mainly allocated to research, start-up activities and phase-out activities, most of which would not be relevant to another state implementing an OPDUR program. For this reason, the analysis to a large extent was limited to costs incurred during the time period June 1994 - June 1997.

Nevertheless, some start-up activities would be relevant to another state implementing OPDUR, and the costs of these activities therefore should be included in the analysis. Dr. Chrischilles identified certain start-up costs that any state implementing an OPDUR program would likely incur, including initial programming and customization costs. Time spent by key staff members on these tasks was considered to be necessary for an enhanced program. In Iowa, these start-up activities eventually benefitted four states (including Iowa) because the OPDUR vendor used the customized product in three other states's OPDUR programs. For this reason, only a quarter of the start-up costs incurred were allocated to the Iowa OPDUR program.

---

[61] Most long-distance calls were made to Unisys, the state's fiscal intermediary. The enhanced OPDUR system would be expected to require many more of these calls because of the more sophisticated programming and management required to maintain the system.

*Adjusting for the Differential Timing of Costs*

In most economic evaluations, costs incurred at different points in time are adjusted to account for the idea that costs incurred later are preferable to costs incurred immediately. This adjustment process is called discounting [Weisbrod (1994)]. Discounting is particularly important when comparing two programs that incur costs at different points in time.

The Iowa OPDUR program was not explicitly compared to another program in this analysis. The costs of OPDUR therefore were not discounted to reflect differential timing. In effect, the implicit comparison in this study was between OPDUR and the computerized systems currently used in most pharmacies.

## Results: Operational Costs of the Iowa OPDUR Program

This section describes important line items from the state, the Iowa Pharmacists Association and the University of Iowa portions of the budgets. Using the methods described above, an operational cost was calculated for every budget line item in each relevant year. The operational costs are summarized below.

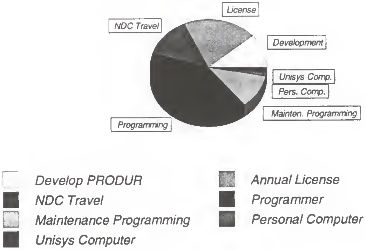
*Iowa*

Initially, most OPDUR operational costs at the state level were related to the integration, testing, modification and documentation of ProDUR™, the OPDUR software received from NDC. The annual license fee for ProDUR™ was $17,500. Additionally, the state was responsible for paying $11,520 in travel expenses for NDC personnel who came to Iowa to implement ProDUR™ on the existing Unisys computer system.

The ProDUR™ software required fairly extensive modifications, and Unisys and University of Iowa researchers made most of these modifications. These tasks required considerable programmer time at the state; in fact, in every year of the demonstration, programming costs were the largest component of state-level operational costs. Other operational costs incurred by the state included initial purchases of computers (one personal computer and one larger computer for Unisys), and programming to maintain the OPDUR system. Exhibit 5.13.1 below displays operations costs for the basic OPDUR model in the second year of the demonstration, when the OPDUR system was first implemented.

*Exhibit 5.13.1: Operations Costs at the State Level*

*Basic OPDUR Model, 1993-1994*



## Iowa Pharmacists Association

The Iowa Pharmacists Association (IPA) was mainly involved in research-related, rather than operational, aspects of the demonstration. The IPA was responsible for coordinating the state liaison board conference calls and meetings. Because the state liaison board dealt with both research and operational issues, half of the funds spent for state liaison meetings were considered to be operational costs. Otherwise, no other costs at the IPA level were included in this analysis.

### *University of Iowa*

The University of Iowa incurred most of the costs of operating OPDUR. In the first year of the demonstration, two key staff members (Dr. Elizabeth Chrischilles and Ms. Criss Jensen) performed initial research and programming that were considered necessary to operate an enhanced OPDUR system. For the enhanced system, therefore, a portion of their time, which cost about $3101 (including fringe benefits), was included as operations costs in the first year of the demonstration. Otherwise, all costs in the first year of the demonstration were related to research activities.

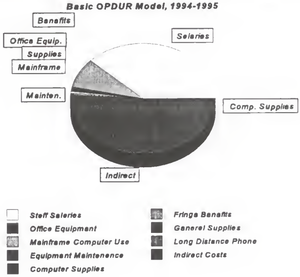
OPDUR was in operation by the last few months of the second year of the demonstration. Most costs related to operations during these months were staff costs. OPDUR demonstration staff were involved in general oversight of operations, database management, and system evaluation, testing and updates. Smaller costs in this year included office equipment purchases, mainframe computer time, long distance phone calls, and office supplies.

The majority of OPDUR operations costs were incurred in years three and four of the demonstration. Indirect costs, staff salaries and fringe benefits comprised the majority of operations costs. Major tasks performed by OPDUR staff members during these years included project coordination, processing of clinical database updates, continued testing and evaluation, and maintenance of software. Other costs included computer and other supplies, office equipment, mainframe computer time, and long distance phone calls.

During the final year of the demonstration, a few months were used to phase-out the OPDUR system. These months were excluded from the analysis since this issue would not be relevant to a state implementing OPDUR. As in the other years, costs incurred in the remainder of the final year were primarily staff salaries and fringe benefits. Small amounts of funds were spent on supplies and mainframe computer time.

Exhibit 5.13.2 displays operation costs incurred at the state level for the basic OPDUR model during the third year of the demonstration. The highest proportion of funds is allocated to indirect costs in these years of the demonstration. Because of the magnitude of indirect costs, the assumptions used to generate the indirect costs estimates are examined in a later section. Salaries and fringe benefits also were important components of operations costs. Other costs, such as supplies, equipment and mainframe costs, were much smaller percentages of total operational cost.

**Exhibit 5.13.2: Operations Costs at the University Level**



## Results: Average Cost Ratios

The average cost per year of operating the enhanced OPDUR program was $199,300.70. For the basic OPDUR program, this yearly, average cost was $158,848.60. In Exhibit 5.13.3 below, the enhanced and basic cost of OPDUR is displayed for every year of operation.

**Exhibit 5.13.3**

**HCFA DUR Demonstration Evaluation**
**Costs Per Year of a Basic and an Enhanced OPDUR System**

| Time Period | Demonstration Year | Operations Costs per Time Period of a Basic OPDUR System | Operations Costs per Time Period of an Enhanced OPDUR System |
|---|---|---|---|
| 6/94 - 9/94 | 2 | $42,092.26 | $52,364.96 |
| 10/94 - 9/95 | 3 | $203,095.90 | $249,947.60 |
| 10/95 - 9/96 | 4 | $142,905.70 | $187,300.60 |
| 10/96 - 6/97 | 5 | $70,926.22 | $97,156.58 |

**Source: Iowa OPDUR demonstration budgets**

The enhanced system, unsurprisingly, is somewhat more costly than the basic system in every year of operation. The extra costs are largely the result of the extra resources (mostly staff resources) required to maintain, test, and update a state-of-the-art OPDUR system. There was reduction in basic and enhanced OPDUR costs from year three to year four. This reduction can be attributed to lower programming costs and the end of the state liaison board meetings by the fourth year of the demonstration.

A large portion of the costs of OPDUR are fixed costs. Regardless of the number of prescriptions to be screened, the OPDUR system must be customized to the state's specifications, tested, and continually maintained and updated. The marginal cost of screening an additional prescription is likely to be close to zero.[62]

During the time period 6/13/94 - 3/31/97, the OPDUR system screened 9,794,987 Medicaid claims, including paid, rejected and reversed claims. This analysis focused on the time period 6/94 - 6/97, when the OPDUR system was in operation. The number of claims screened between 6/94 - 6/97 was extrapolated based on the number of claims screened between 6/94 - 3/97. It was estimated that in total, 10,685,440 claims were screened by OPDUR during its three years of operations. Exhibit 5.13.4 below displays the average cost per prescription screened for the basic and enhanced OPDUR programs.

---

[62] From the perspective of the pharmacist's time , the marginal cost may not be zero, but it is likely to be small. In any case, pharmacists' time costs were not included in this analysis.

**Exhibit 5.13.4**

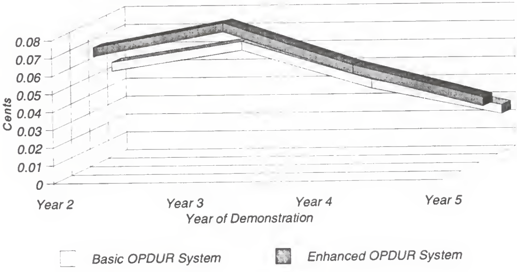
**HCFA DUR Demonstration Evaluation**
**Average Cost Per Prescription Screened by OPDUR**

| | Basic OPDUR System | Enhanced OPDUR System |
|---|---|---|
| average cost per prescription | $.043 | $.055 |

Sources: Iowa OPDUR demonstration budgets

On average, the enhanced and basic OPDUR programs cost approximately four to six cents per prescription screened. Over time, one would expect the average cost per prescription to decrease in both a basic and an enhanced OPDUR program. Initially, there might be implementation problems and programming costs that would diminish in importance as the OPDUR system matures. Exhibit 5.13.5 confirms this hypothesis — as the demonstration moved into its final years, the operating costs of OPDUR declined and the number of prescriptions screened increased, leading to a considerable drop in the cost of OPDUR per prescription screened.

*Exhibit 5.13.5: Cost of OPDUR Per Prescription Screened*



## Sensitivity Analysis

Like any cost analysis, this study was based on many assumptions and estimates that are subject to some degree of uncertainty. To examine the sensitivity of the results to this uncertainty, potentially controversial issues were identified and the calculations were re-worked incorporating a range of reasonable values. If the results do not change considerably, the sensitivity analysis suggests that the results are fairly robust. Otherwise, if different, reasonable assumptions and estimates do change the results significantly, it is important to understand why these changes are occurring, and what implications these changes have for other states using the analysis to implement OPDUR programs.

In this section, the method used to calculate indirect costs, and assumptions made about the interest rate and the length of useful life of capital were subjected to sensitivity analyses.

***Indirect Costs Calculation.*** Indirect costs were calculated by estimating the proportion of direct, operations (rather than research), costs in each year and then by multiplying this proportion by the funds actually spent on indirect costs in that year. It was argued that this method was reasonable because indirect costs are assumed to be proportional to direct costs, and research-related indirect costs were excluded from the analysis.

This method, however, is based on the indirect costs figures calculated using the University of Iowa rate schedule. The indirect cost rate charged by a university may not be a good approximation for the indirect costs that a state implementing OPDUR would incur. Furthermore, indirect costs comprised one of the largest budget line items in every year, making it an important part of the analysis to test for robustness.[63]

Soumerai and Avorn (1986) estimate indirect costs to be 20 percent of non-contractual, direct costs for academic detailing programs based in medical schools. This method was used as an alternative; two line items that represented contractual payments to NDC (annual license fee and travel for NDC staff) were excluded from direct operating costs, and 20 percent of the remaining direct, operating costs was used as an estimate for indirect costs. As displayed below in Exhibit 5.13.6, this alternative method of calculating indirect costs changed the average cost of OPDUR per prescription only slightly. For both the basic and enhanced OPDUR systems, the cost per prescription was higher using the alternative method of calculating indirect costs.

**Exhibit 5.13.6**

**HCFA DUR Demonstration Evaluation**
**Average Cost Per Prescription Screened by OPDUR:**
**Calculated Using Two Different Methods of Estimating Indirect Costs**

| | **Original Method** | | **Alternative Method** | |
|---|---|---|---|---|
| | **Basic** | **Enhanced** | **Basic** | **Enhanced** |
| average cost per prescription | $.043 | $.055 | $.047 | $.061 |

*Sources:* Iowa OPDUR demonstration budgets

*Capital Costs.* Capital purchases, including computers and other office equipment, were annuitized using a five percent interest rate and an estimated useful life of five years for the computers, and ten years for the office equipment. These assumptions about useful life were conservative. In particular, it is possible that the computers would not need to be replaced after five years. The analysis was reworked assuming that the computers have a useful life of seven years and the office equipment has a useful life of 15 years. With these new assumptions, the average cost per prescription for the basic and the enhanced systems was virtually unchanged.

---

[63] Indirect costs were 24-25 percent of total costs summed over the years OPDUR was in operation.

The Department of Health and Human Services has recommended that a three percent interest rate be used when evaluating public sector projects [DHHS (1996)]. Again, when the office equipment and computers were annuitized with a three percent discount rate, using ten and five years as estimates of useful life, the average cost per prescription was virtually unchanged. The least conservative combination (a three percent discount rate, a useful life of seven years for the computers, and a useful life of 15 years for the office equipment) also yielded an average cost per prescription that was almost the same as the one that resulted from the most conservative estimates (a five percent discount rate, a useful life of five years for the computers, and a useful life of ten years for the office equipment). This analysis therefore is quite robust to different assumptions about discount rates and useful life of capital.

### The Costs of OPDUR Relative to Other Methods of Improving Prescribing

Although the effectiveness of OPDUR systems has not been addressed until now, there is a growing body of literature on the effectiveness of other methods of improving prescribing. The literature includes randomized, controlled studies of academic detailing,[64] educational mailings, group educational sessions, retrospective drug utilization review (RDUR) feedback letters,[65] payment restrictions and clinical pharmacist systems. These studies indicate that some interventions, including academic detailing and feedback letters, have the potential to change prescribing practices.

Nevertheless, in this body of work, there is almost no mention at all, of the costs of implementing these interventions. Some authors [including Smalley et. al. (1995)] and Ahluwalia et. al. (1996)] state that the intervention under study is not costly, but they make no attempt support this assertion with data. One exception is Stergachis et. al. (1987), who in their assessment of clinical pharmacy services acknowledge that the cost savings associated with the program were smaller than the program's operating costs. These researchers estimate that while the clinical pharmacy services program was associated with a drug cost savings of $.45 per enrollee per year, the operating costs of the intervention were about $1.30 per enrollee per year [Stergachis (1987)].

Avorn and Soumerai (1983) and Soumerai and Avorn (1986) also consider the costs of academic detailing. These researchers collected data on the cost of supplies, training, managerial time, and the time spent by pharmaceutical educators. They report that for a half-year program targeting 10,000 physicians, the academic detailing program would yield net savings between $67 and $111 per physician (1980 dollars) to Medicaid, depending on the degree of substitution to OTC drugs [Soumerai and Avorn (1986)]. The medications targeted by this program represent only 1.7 percent of all drugs prescribed in office-based practices, suggesting that programs with broader scopes may lead to even greater cost savings.

---

[64] Academic detailing is defined as face-to face, informational discussions between trained counselors and prescribers.

[65] Please see Chapter Six for evaluations of retrospective drug utilization review letter interventions in Iowa and Maryland.

Most recently, Collins et. al. (1997), evaluate a retrospective drug utilization review letter intervention in Wisconsin. These researchers find that the program led to drug cost savings of $30, 434 in one year. They estimate that preparation and mailing of the letters cost $6272 [Collins et. al. (1997)].

Ideally, one would like to compare the costs and benefits of OPDUR to the costs and benefits of other methods used to improve prescribing. For example, if cost-effectiveness ratios were available for academic detailing, RDUR feedback letters, and OPDUR, a comparison could be made between these three different approaches. Unfortunately, there is little if any information on the costs of these methods; research efforts to link the interventions to patient outcomes are also limited. At this time, therefore, it is not possible to evaluate the cost-effectiveness of OPDUR to other interventions that share similar goals. It is hoped that this analysis will encourage other researchers to develop cost estimates of interventions in this area.

## Limitations of the Analysis and Conclusions

This analysis is subject to several limitations. First, all information was collected from OPDUR demonstration budgets, which combine the costs of research and operations activities. Although this analysis attempted to extract the operational costs from each line item using estimates from the demonstration's principal investigator, there is considerable uncertainty about these estimates. Furthermore, although two methods were used to estimate indirect costs, both methods are based on a university indirect cost rate, which may not be similar to the indirect cost rate that would be incurred by a state Medicaid program.

Costs not included explicitly in the budgets, such as time costs, were assumed to be small in magnitude. Consequently, no attempts were made to measure intangible costs. The lack of estimates for these costs may have resulted in an under-estimation of OPDUR costs. Ideally, if the cost data had been collected as part of a process evaluation, the importance of time costs and other intangibles could have been assessed. This kind of data also may have allowed a more precise separation of operations costs from research costs.

The limitations of the data available for this study affect the interpretation and the generalizability of the results. Costs were estimated for a basic and an enhanced OPDUR model. The basic model was intended to represent an OPDUR system similar to the ones currently in place in most states, while the enhanced model was a state-of-the-art system, developed in part for research and evaluation purposes. The costs of both models were estimated using information on an OPDUR demonstration that was based at a university. All of the University of Iowa line items (which constitute the majority of the budget in every year of the demonstration) therefore reflect the unique structure and costs of a university.

Consequently, the OPDUR cost estimates presented in this chapter should be interpreted as estimates of a university-based basic OPDUR system and a university-based enhanced OPDUR system. These cost estimates may be particularly useful to a state that is planning to implement an OPDUR system in conjunction with a university. Although the estimates are subject to the limitations discussed above, these types of OPDUR programs may face costs that are very similar to those incurred by the Iowa OPDUR demonstration, especially if a considerable amount of programming, management and technical support are provided by a university.

Medicaid programs with plans for fully state-run OPDUR programs perhaps should use the estimates with greater caution. While a fully state-run program may face many of the same types of costs as an OPDUR system jointly implemented by the state and a university, the magnitude and structure of the costs may be quite different. For example, while a state might hire a full-time programmer to maintain the OPDUR system, a university might rely on a series of part-time graduate students for programming support. These types of differences most likely will affect the costs of the OPDUR system.

# 6.0 Retrospective Drug Use Review

## 6.1 Introduction

As mandated by the 1990 Omnibus Budget Reconciliation Agreement (OBRA 90), all state Medicaid programs currently operate retrospective drug utilization review (RDUR). RDUR, the focus of this chapter, is an authorized, structured, and continuing program that reviews, analyzes, and interprets patterns of drug use (after the drug use has occurred) in relation to predetermined criteria and attempts to reduce inappropriate drug prescribing.[1] State Medicaid programs commonly implement RDUR by sending educational feedback letters to targeted physicians and pharmacists. These letters usually inform physicians and/or pharmacists of their inappropriate use of a particular drug with a particular patient or set of patients. Although RDUR letters generally list the name(s) of patient(s) found to have prescription problems, the intent of the intervention is to effect lasting, broad improvements in prescribing quality that extend beyond the patients named in the letters.

RDUR occurs over long periods of time, letters typically reach physicians months after the events that precipitated them. By necessity, RDUR letters tend to be imprecisely targeted. In the case of Medicaid DUR, the DUR board typically only has access to the information captured by the patient's Medicaid claims. Furthermore, because programs are designed on limited budgets, they often maximize the use of computers and minimize professional review; the computer algorithms are by necessity crude and lead to many false positives. This computerized screening is often conducted by outside vendors using existing algorithms with the result that DUR boards do not have much control over it.[2]

Despite these limitations, findings from recent, published evaluations suggest that RDUR letter interventions are effective in changing the prescriptions of those patients named in the letters. Rascati et. al. (1996), Zimmerman et. al. (1994), Guo et. al. (1995), and Collins et. al. (1997) all find that Medicaid RDUR letters have at least some direct impact on the prescriptions of patients who were named in the letters.[3] None of these studies, however, assess the broader impact of the RDUR letters that extends to patients whose prescribers received letters but who were not named in the letters. In this chapter, benefits of the letter intervention that extend to targeted prescribers' patients who are at-risk but who were not named in the letters are termed "spillover effects."

This chapter includes analyses of the direct and spillover impact of three Medicaid letter interventions in Iowa and Maryland: (1) a 1994 effort to prevent ulcers by encouraging prescriptions of misoprostol for elderly patients taking non-steroidal anti-inflammatory drugs (IA); (2) a 1995 effort to promote appropriate asthma treatment by increasing the use of short-acting bronchodilator among patients taking salmeterol (IA); and (3) a 1994 effort to reduce the inappropriate long-term use of Zantac, an anti-ulcer medication (MD). All three interventions were assessed using quasi-experimental methods. As described in detail in the following sections, a comparison group is identified for each intervention that is comprised of Medicaid patients whose prescription profiles would have triggered an intervention letter, but whose profiles were not screened because of timing, manpower limitations and other issues.

---

[1] The definition draws on two similar definitions one appearing in: Lipton, Helene Levens, "Drug Utilization Review in Ambulatory Settings: State of the Science and Directions for Outcomes Research," *Medical Care*, 31(7): 1069-1082; and Lipton, Helene L, and Soumerai, Stephen B., "Computer-Based Drug Utilization Review - Risk, Benefit, or Boondoggle " *New England Journal of Medicine*, 332(24): 1641-45.

[2] A broader analysis of problems with RDUR might include the lack of agreement over appropriate therapeutic criteria for screening, the limits on the types of problems that RDUR can address, and the fact that the crudeness of screening techniques tends to yield a high proportion of false negatives as well as false positives.

[3] These studies are described in more detail in Chapter One of this report.

To learn whether or not the three letter interventions were effective in making lasting changes in prescriber behavior, treatment and control group patients' Medicaid drug claims for the targeted drugs were compared after the intervention. Additionally, Medicaid drug claims of other at-risk patients of treatment prescribers were examined to determine whether the interventions had spillover effects beyond the patients named in the letters. Statistical methods were used to control for confounding variables and time trends.

The findings suggest that all three interventions were effective in improving the quality of prescriptions written for patients named in the intervention letters. After the Iowa misoprostol letter intervention, eight percent of targeted patients were using misoprostol compared to one percent of control group patients. The Iowa salmeterol letter intervention led to an eight-point increase in the percentage of patients using both salmeterol and a short-acting bronchodilator in the follow-up period. In Maryland, the Zantac letter intervention resulted in a large reduction in Zantac use; three months after the intervention, 44 percent of treatment group patients were still using Zantac compared to 73 percent of control group patients. These results are consistent with the small body of research in this area. RDUR letters are effective in a narrow sense — they can improve prescribing quality specifically for patients named in the letters.

There was only limited evidence, however, that RDUR letter interventions have spillover effects. The benefits of the Iowa letter interventions did not extend beyond patients who were named in the letters. Although there was some evidence of spillover effects in the evaluation of the Maryland letter intervention, it was not clear how limitations in the study design for that analysis affected the results.

Overall, then, the findings of the three evaluations suggest that Medicaid RDUR letter interventions are effective in improving the quality of prescriptions for patients named in intervention letters, but not necessarily for other patients of targeted prescribers. These studies indicate that, when used alone, RDUR letter interventions do not promote lasting, broad changes in prescribing practices. Future research should continue to examine the spill-over effects of RDUR interventions as well assess the costs of these interventions. This information will be useful to policy makers as they implement strategies to improve the overall quality of prescriptions in the Medicaid population.

## 6.2 The Impact of RDUR — Iowa Medicaid DUR Board Letters on Misoprostol

### Introduction

This study examined the impact of RDUR letters mailed to physicians and pharmacies concerning patients who could potentially benefit from adding misoprostol to their drug regimens. Misoprostol is the one drug approved by the FDA for the prevention of gastric ulcers induced by non-steroidal anti-inflammatory drugs (NSAIDs). While the emphasis was certainly prophylaxis via misoprostol, the educational materials that accompanied the letters also mentioned that non-acetylated salicylates carry a lower risk of ulceration than other NSAIDs.

Physicians and pharmacies were selected to receive these letters based on the review of a sample of eligible patients' profiles. Because only a sample of eligible profiles were reviewed, this study was able to create a comparison group of physicians in the same state (Iowa) during the same time period (1994-1995) who did not receive letters but who would have if all eligible patient profiles had been reviewed. The letters' impacts were then estimated by comparing patients of treatment and comparison prescribers.

This study found that the letters led to a statistically significant increase in the utilization of misoprostol among the patients named in the letters; eight percent of these patients received misoprostol as opposed to one percent of the comparison group. It found no change in the use of misoprostol among two other populations of patients of contacted prescribers who appeared to be good candidates for misoprostol but were not actually named in the letters. The study did not find any evidence that the letters led to an increase in the utilization of non-acetylated salicylates among patients named in letters; it did find some rather questionable evidence of increased use of non-acetylated salicyclates among one of the spillover populations. Moreover, the study found no evidence that the letters caused reductions in the percentage of patients at risk in the follow-up period either among those patients named in the letters or among either of the two spillover populations.

### Previous Research and Contribution of This Study

Three other recently published, peer-reviewed studies examined the effectiveness of Medicaid RDUR letters using an experimental or quasi-experimental design. Zimmerman et al. (1994) studied a 1988 Wisconsin intervention designed to reduce the prevalence of excessive duration of full-dose histamine-2 receptor antagonists (H2RAs) using a quasi-experimental design. In this intervention, a physician received a letter if she had three patients who were exceptional according to Wisconsin's therapeutic criteria; the exceptional patients of these physicians formed the treatment group while exceptional patients of physicians with two or fewer such patients formed the comparison group. This study observed a greater reduction of H2RA use in the intervention group than in the comparison group and a reduction in drug expenditures which exceeded the cost of targeting and notifying prescribers.

Guo et al. (1995) evaluated a 1993 Alabama Medicaid intervention aimed at reducing the costs of six anti-ulcer agents. The study group consisted of prescribers who received letters concerning therapeutic exceptions and who consistently prescribed the drugs in question while the comparison group contained prescribers who consistently prescribed the drugs in question but who had no exceptional patients according to the screening criteria and received no letters; changes between the pre- and post-intervention period for the study group scaled by the corresponding changes for the comparison group were analyzed. This study found that prescribers' behavior changed after the DUR letters were sent out resulting in a reduction in the average marginal days of drug therapy and drug reimbursement per recipient.

Finally, Rascati et al. (1996) assessed the effect of Texas Medicaid RDUR letters indicating duplicative anti-ulcer medications. This study was based on a true random experiment in that patients whose profiles indicated the problem were divided randomly into two groups; letters were sent to the physicians of patients in one group; patients in the other group served as controls. Profiles generated six months after the letters were sent indicated that 48 percent of the patients in the experimental group were still receiving concurrent therapy compared with 64 percent of patients in the control group.

The current study adds to this body of evidence that RDUR letters can be an effective way to change prescribing. Although the Iowa misoprostol intervention was not designed with the explicit goal of offering a random experiment for evaluation, the way that it was implemented offered the evaluator a *de facto* random experiment. This is the cleanest and most credible evaluation design.

This study differed from the three described above in that the misoprostol intervention was intended to promote the prescribing of a target drug, misoprostol, while, in each case described above, the intervention was intended to decrease utilization and/or costs of target drugs. Thus, this study adds robustness to the conclusion that RDUR can be effective by demonstrating that it can encourage as well as discourage drug use. Many believe that the underutilization of potentially effective drugs is as significant a therapeutic and cost issue as the over utilization of unnecessary or potentially harmful drugs. This study showed that, at least in one case, underutilization could be effectively addressed via RDUR.

Finally, this study expands on previous work by distinguishing between two groups of patients that could be affected by RDUR, the patients whose profiles trigger the letters and other patients of targeted prescribers, and by separately assessing effects on each group. Often an implicit or explicit goal of RDUR is lasting change in a prescriber's behavior; the patients named in letters are examples of the problem, but the educational impact of the letters is hoped to extend to other current and future patients of targeted prescribers as well. After assessing the direct impact, this study also examines whether this broader impact can also be observed.

### *Misoprostol*

Fifteen percent of all prescriptions in the United States are written for NSAIDs, and gastropathy caused by NSAIDs is the most common drug-induced adverse effect occurring in the United States. Gastrointestinal (GI) complications in rheumatoid arthritis patients alone result in 2,600 deaths and 24,000 hospitalizations each year. Moreover, the cost of the management of GI side effects accounts for 31 percent of the total treatment costs for arthritic patients.

NSAID-induced ulcers are characterized as silent ulcers; 60 percent of patients who have a severe GI bleed due to NSAID use have no symptoms prior to the bleed. For this reason, physicians must rely on risk factors to determine which patients require prophylactic measures; age, prior GI disease, use of corticosteroids, use of alcohol, and use of tobacco are all risk factors. Two other factors, anticoagulant therapy and high risk of surgical complications, increase the risk of debilitating complications if GI bleeding does occur. The medical literature suggests that the presence of two or three factors places a patient in a high-risk group.

Among high risk patients, there are several ways to reduce the risk of gastropathy. One option is to replace the NSAID with an alternative drug such as acetaminophen, analgesic creams, or a non-acetylated salicylate. Another is to use the lowest effective NSAID dose and to avoid concurrent use of multiple NSAIDs. Finally, the FDA has approved one drug, misoprostol, for the prophylaxis of NSAID-induced gastric ulcers. Common treatment options for peptic ulcer disease, including H2RAs, antacids, and sucralfate, have not been shown to be effective in the prevention of NSAID-induced gastric ulcers.

Diarrhea is a common side effect of misoprostol leading some patients to resist taking the drug, especially when they are asymptomatic.

### The Iowa Misoprostol Intervention

The Iowa Medicaid DUR Board chose to focus some of its RDUR efforts on NSAID-induced ulcers because, as emphasized above, they were a common adverse drug effect, because there were clear recommendations for managing drug therapy to minimize risk, and because profile review suggested that many patients who might benefit from misoprostol did not receive it. Profile review also suggested that some patients might be receiving ineffective drugs as prophylaxis for the gastric ulcers caused by NSAIDs.

The Board developed educational materials on the prophylaxis of NSAID-induced gastric ulcers and on misoprostol in particular. (See Appendix K) These formed the basis of an article on these issues in the Summer 1994 issue of the *DUR Digest*, a quarterly publication mailed to all physicians and pharmacies participating in Iowa Medicaid and operating in Iowa. At the same time, the Board decided to conduct a focused RDUR intervention designed to encourage the appropriate use of misoprostol.[4]

The process that selected the providers who ultimately received letters was rather complex. The first stage was a computerized review of Iowa Medicaid drug claims with dates of service between January and June of 1994 that identified all Medicaid recipients over age 60 who received both an NSAID (AHFS class 280804) and a maintenance dose of a GI drug (AHFS class 564000) during that time period. This review identified 3,103 Medicaid recipients in a broad risk group for gastropathy due to NSAID use, age, and possible use of an inappropriate prophylaxis.

The second stage was a manual review of patient drug profiles by the professional staff of the Iowa Medicaid DUR Board. Because of limited person power and the large volume of cases identified via computerized review, only 747 of the 3,103 recipients identified by the computerized review had their profiles extracted for manual review. This sample was randomly selected. Although this review involved professional judgement, certain general principles were applied. (1) A patient was considered to be at risk if that patient exhibited concurrent long-term NSAID use and use of a GI drug at a maintenance level.[5] The maintenance dose was interpreted as an indication of possible intended prophylaxis. (2) Patients also were selected for the intervention if it was deemed that their providers might benefit from education. Examples included (A) patients with NSAID usage followed by a GI drug indicating the possibility of an NSAID-induced ulcer and (B) patients with concurrent NSAID and GI drug use with both drugs starting on the same day and the GI drug at a maintenance, rather than a therapeutic, dose, again possibly indicating that the GI drug might have been intended as prophylaxis.

---

[4] The net influence of the letters may have been attenuated by the earlier release of educational information via the *DUR Digest*.

However, patients were excluded from the risk group if (1) the NSAID was once-daily aspirin, (2) the NSAID was a non-acetylated salicylate, which has a very different mechanism of action than most NSAIDs and places that patient at a lower risk for a GI bleed, or (3) misoprostol was used. A total of 270 patients were considered to be at risk based on these criteria.[6]

A note was also made if a profile featured any additional risk factor: (1) continuous anticoagulant therapy, (2) use of two or more concurrent NSAIDs, or (3) continuous corticosteroid therapy. Of the 270 at-risk profiles, 31 also indicated an additional risk factor.

The Iowa Medicaid DUR Board had a dual purpose in their intervention; they sought both to improve therapy for the individual patients whose profiles were examined and to influence prescribers' behavior more generally. In keeping with these goals, they sent letters to all prescribers who were treating either one patient who met the selection criteria and exhibited an additional risk factor placing him at a higher risk for an NSAID-induced ulcer or two or more patients meeting the selection criteria. A total of 73 prescribers received letters. Letters were also sent to pharmacies that filled NSAID or GI prescriptions for either an at-risk patient with an additional risk factor or two or more at-risk patients. A total of 83 pharmacies received letters. See Appendix K for copies of the letters, which were mailed January 16, 1995.

It should be noted that this intervention was one of the Iowa Medicaid DUR Board's earliest problem-focused studies. RDUR in Iowa has evolved considerably between 1994/1995 and the time of this writing. This intervention was selected for study because, at the time the study began, 1995 was the most recent year of data available to Abt Associates.

### Research Method

#### *Study Populations*

This study posed four questions, each involving a different study population:

---

[5] The guidelines developed for manual screening were used to create a computer screening algorithm for this study. The parenthetical numbers and letters in these paragraphs are intended to help the reader understand the logic of that algorithm, discussed below.

[6] Two of the patients were later excluded because the prescriber had died or could not be identified.

1. *Did the RDUR intervention lead to changes in drug therapy for the patients specifically identified in the letters?* This is the most straightforward potential effect of the letters.

2. *In particular, did the RDUR intervention lead to changes in drug therapy for the patients identified in the letters who* ***continued*** *to receive NSAIDs or GI drugs from the prescribers who received the letters?* Although letters would be intended to affect all named patients, one would expect their effect to be concentrated among the subgroup of patients who continued to see the same doctor.

3. *Did the prescribers who received RDUR letters alter their prescribing for additional patients beyond those named in the letters? Specifically, did the letters lead to changes in drug therapy for other Medicaid patients who were at risk during the review period?* A stated goal of the intervention was to change prescribers' behavior beyond the patients identified in the letters. One group for which this effect might exist would be patients of contacted prescribers who were at risk during the review period but were not named in the letters because their profiles had not been included in the random sample that received manual review.

4. *Did the RDUR letters lead to changes in drug therapy for other Medicaid patients of the contacted prescribers who were* ***newly at risk*** *in the period immediately following the letters?* One might expect the letter to affect prescribing for patients who were new candidates for misoprostol in the follow-up period as well as for those who were candidates in the review period. This question is examined by using all patients in the broad risk group (over 60 and receiving both an NSAID and a maintenance dose of a GI drug) in the follow-up period who were not in that group during the review period.

### Outcomes and Hypotheses

The intervention was intended to promote the utilization of therapy shown to be effective for the prophylaxis of NSAID-induced ulcers and to reduce the numbers of patients at risk for these ulcers. This study defined seven possible outcomes for a patient deemed to be at risk during the review period.

1. *Patient using misoprostol in the follow-up period; change made before letters were mailed.* An at-risk patient might begin misoprostol before January 16, 1995, the date the RDUR letters were mailed. This change could not be attributed to the RDUR letters though it might result in part from the educational efforts of the Iowa Medicaid DUR Board via their Summer 1994 article in the *DUR Digest.*

2. *Patient using misoprostol in the follow-up period; change made after letters were mailed.* A patient might begin misoprostol after the letters were mailed and have claims during the follow-up period (March 1- August 31, 1995). The RDUR letters emphasized prophylactic treatment and increases in the use of misoprostol were the main expected effect of the intervention.

3. *Patient using a non-acetylated salicylate in the follow-up period; change made before letters were mailed.* A patient might replace his original NSAID with a non-acetylated salicylate before January 16, 1995. Again, this change, thought to reduce the risk of an

NSAID-induced ulcer, could not be attributed to the RDUR letters but might be due in part to the other educational efforts of the DUR Board.

4. *Patient using a non-acetylated salicylate in the follow-up period; change made after letters were mailed.* A patient might replace his original NSAID with non-acetylated salicylate during the follow-up period. The educational materials that accompanied the RDUR letters mentioned that non-acetylated salicylates are less likely to cause gastric ulcers than other NSAIDs. There was a weaker expectation that the intervention would increase the likelihood of this outcome.

5. *At risk.* A patient might remain at risk as defined by the criteria above. The intervention was expected to reduce this outcome.

6. *Not at risk.* A patient might leave the defined risk group by ceasing GI drugs or NSAIDs altogether or by altering his drug utilization. In principle, the intervention could increase the likelihood of this outcome by raising the probability that a patient would stop NSAIDs or the GI drug, if it was used solely for prophylaxis. However, such an effect would be very attenuated so that the practical expectation would be that the intervention would not affect this outcome.

7. *No longer eligible for Medicaid.* A patient might lose his eligibility for Medicaid before the start of the follow-up period; this could not be attributed to the intervention. In that case, if the patient continued to see the same physician, then the RDUR letters might still affect therapy. However, because the current study relies on Medicaid data, these outcomes cannot be counted specifically.

### Analysis Design

In researching the impact of the RDUR intervention, one would like to compare what happened given the RDUR intervention and what would have happened if the RDUR intervention had never occurred. Unfortunately, this comparison is not literally possible. The best feasible comparison is to find a suitable comparison group and assume that the experience of that group mirrors what would have happened absent RDUR.

The design for this study took advantage of the fact that only a random sample of the patient profiles identified via computerized review actually underwent manual screening. The study defined treatment prescribers as those who received letters; it defined comparison prescribers as those who would have received letters if all profiles had been screened but, in fact, did not receive letters because the random sample did not include all of their at-risk patients. Treatment and comparison groups of patients were then defined based on their relationships with the respective groups of prescribers.

In the analysis of Question 1, the direct effect of the letters, outcomes for the patients who were actually named in letters were compared to outcomes for at-risk patients of comparison prescribers. The analysis of Question 2 used similar treatment and comparison groups but restricted the sample to patients who continued to receive either NSAIDs or GI drugs in the follow-up period from the same physician who prescribed them in the review period.

A prescriber who received a letter might also have had several patients who were at risk in the review period but not in the random sample and therefore not named in the RDUR letters. These patients formed the treatment group for Question 3, the first of two questions concerning spillover effects onto other patients. The comparison group was again the at-risk patients of comparison prescribers.

Finally, the treatment group for Question 4 was patients of treatment prescribers who were new members of the broad risk category. Specifically, the treatment group for this question consisted of patients who were over 60 and received NSAIDs and GI drugs in the follow-up period but did not meet these criteria during the initial review period. The comparison group was the corresponding patients of comparison prescribers.

Exhibit 6.2.1 is a schematic representation of this analysis design.

**Exhibit 6.2.1**

**HCFA DUR Demonstration Evaluation**
**Impact of RDUR Letters Regarding Misoprostol Analysis Design**

| | | | | |
|---|---|---|---|---|
| patients at risk in review period (narrow definition) | prescriber would have received letter if all profiles had been screened | prescriber received letter | pt in RS/letter named this pt | Question 1,2 treatment |
| | | | pt not in RS/letter did not name this pt | Question 3 treatment |
| | | prescriber did not receive letter | | Question 1,2,3 comparison |
| pt member of broad risk group in follow-up period, not member in review period | prescriber would have received letter if all profiles had been screened | prescriber received letter | | Question 4 treatment |
| | | prescriber did not receive letter | | Question 4 comparison |

Although the design of this study is organized around physicians, both physicians and pharmacists received letters and either party might have responded to a letter by taking steps to improve drug therapy. Note that any effect estimated in response to Question 1, that is a difference in therapy observed for patients named in letters relative to comparison patients, could have resulted either from physician or pharmacist action. Questions 2, 3, and 4 are designed to zero in on effects that work through physician; associated estimates may be biased toward zero to the extent that pharmacist effects exist.

### Selection of Comparison Group

The Iowa Medicaid DUR Board identified at-risk patients via computer and manual screening and then selected treatment prescribers. The research design just presented required recreating that process and applying it to the entire Medicaid population, i.e. performing "manual screening" for all members of the broad risk group. A computer screening algorithm was developed for this study that mimicked the initial computer screening performed by the Iowa Medicaid DUR Board and then imposed the guidelines developed over the course of manual screening and described above. This algorithm was able to correctly reclassify 82 percent of the profiles originally selected for manual screening.[7] The exact results of this validation appear in Exhibit 6.2.2.

**Exhibit 6.2.2**

**HCFA DUR Demonstration Evaluation**
**Impact of RDUR Letters Regarding Misoprostol**
**Validation of Computer Simulation of Manual Review**

| | | Computer Simulation | |
|---|---|---|---|
| | | at risk (247) | not (500) |
| Manual Profile Review | at risk (268) | 192 | 76 |
| | not (481) | 57 | 424 |

*Sources: Iowa Medicaid claims, records of the Iowa Medicaid DUR Board.*

Once the screening algorithm had been validated, it was used to select the treatment group for Question 3 and the comparison groups for Questions 1, 2, and 3. Treatment and comparison groups for question 4 were selected as patients of treatment and comparison prescribers who met the criteria for the broad risk group (over age 60, NSAID use, GI drug use) in the follow-up period but not in the review period.

The final step was to classify outcomes for each patient based on Medicaid claims in the follow-up period. Because letters had been mailed on January 16, 1995, March 1 to September 30, 1995 was selected as the follow-up period. (A six-month follow-up period also meant that the screening algorithm could be used to determine which patients remained at risk and that the results would be comparable to those based on the review period.) Outcomes were classified hierarchically with "misoprostol on or before 1/16/95" taking precedence, followed in order by "misoprostol after 1/16," "non-acetylated salicylate on or before 1/16/95," "non-acetylated salicylate after 1/16," "still at risk," "no longer at risk," and "no longer eligible for Medicaid."[8] Note that individuals are classified based on their Medicaid claims records which may at times be imperfect or incomplete. If this biases the estimates at all, the bias is toward a zero treatment effect.

### Results

The most immediate effect of the letters was hypothesized to be an increase in misoprostol use among patients named in letters. Indeed, in this group, eight percent started misoprostol during the follow-up period. (See the first row of Exhibit 6.2.3.)[9] This can be contrasted with one percent of patients in the comparison group. This seven percentage point difference was statistically significantly different from zero; a 95 percent confidence interval for this difference ranged from 4 to 10 percentage points. Thus, the letters apparently did have the expected direct effect of increasing misoprostol use, although it was still the case that more than 90 percent of patients who were at risk in the review period did not receive misoprostol in the follow-up period.

---

[7] Details of the screening algorithm are available from Abt Associates upon request.

[8] Six percent of the patients at risk during the review period had no claims for which an individual prescriber could be identified. These patients were omitted from the analysis.

If the observed effect did indeed result from the letters, then one would expect it to be especially pronounced among patients who maintained contact during the follow-up period with at least one of the physicians who received letters. The second row of the table indicates that 12 percent of these patients started misoprostol after January 16 and took it during the follow-up period as opposed to two percent of their counterparts in the comparison group. This too is consistent with the letters having the expected direct effect on misoprostol prescribing.

The next two rows of the table show the percentage of patients beginning misoprostol after January 16 in the spillover populations. The first spillover population is those at-risk patients who were not themselves named in letters but whose prescribers did receive personal letters and educational materials. One percent of these patients began misoprostol as did one percent of the comparison population. The second spillover population is patients newly in the broad risk group in the follow-up period whose prescribers received letters. Three percent of these patients began misoprostol as did three percent of comparable patients of comparison prescribers. Thus, although there was clear evidence of a direct effect on misoprostol use, there was no evidence of any spillover effects.

---

[9] The discussion and exhibits are organized around the three outcomes (misoprostol use, non-acetylated salicylate use, and remaining at risk) that the intervention was expected to influence. Appendix K contains information on all seven outcomes organized by study population.

**Exhibit 6.2.3**

**HCFA DUR Demonstration Evaluation**
**Impact of RDUR Letters Regarding Misoprostol**
**Percentage of Patients Receiving Misoprostol During Follow-Up Period (Drug Started After Letters Mailed)**

| | Tr | Comp | Difference[a] | | |
|---|---|---|---|---|---|
| | pct | pct | pct | sig | 95 pct CI |
| Direct effect: patients named in letters | 8% | 1% | 7% | *** | 4% to 10% |
| Direct effect: patients named in letters who continue with same prescriber | 12% | 2% | 10% | *** | 5% to 15% |
| Spillover 1: other "at risk" patients of prescribers receiving letters | 1% | 1% | 0% | NS | -2% to 2% |
| Spillover 2: new patients of prescribers receiving letters | 3% | 3% | 0% | NS | -2% to 2% |

[a] Percentages were rounded for presentation but differences and test statistics were calculated before rounding. The standard test for a difference in proportions was used. * denotes 90 percent confidence that true value of the difference is not zero. ** denotes 95 percent confidence and *** denotes 99 percent confidence. NS indicates a statistically insignificant difference.

*Sources:* Iowa Medicaid claims

A second potential effect of the letters was thought to be an increase in the percentage of patients receiving a non-acetylated salicylate in the follow-up period. As Exhibit 9-4 shows, one percent of patients named in letters had this outcome as did one percent of comparison patients. The difference was statistically insignificant. If one only considers the subset of patients who also continued to receive NSAIDs or GI drugs from a review-period prescriber, then zero percent of the treatment patients began a non-acetylated salicylate while one percent of comparison patients did, again an insignificant difference. Thus, there is no evidence that the letters had the direct effect of increasing the usage of non-acetylated salicylates among patients mentioned.

Surprisingly, there was some debatable evidence of a spillover effect for this outcome. Among at-risk patients who were not named in letters but whose prescribers did receive them, three percent received a non-acetylated salicylate in the follow-up period compared to one percent of the appropriate comparison group. This difference was statistically significant at the 90 percent confidence level (borderline statistically significant.) However, two facts cast doubt on the conclusion that this does indeed represent a spillover effect. First, it seems odd that there would be a spillover effect for non-acetylated salicylate when there was no direct effect. Second, as shown in Appendix K, even before the letters arrived, treatment physicians prescribed non-acetylated salicylates for a greater proportion of their patients who were new members of the broad risk group than did comparison prescribers. This higher underlying rate of utilization does not necessarily imply a higher rate of *changing* to a non-acetylated salicylate, as seen in the first spillover population, but it does raise questions concerning whether these two groups of physicians were completely comparable with regard to this class of drugs.

Within the second spillover population, new members of the broad risk category, five percent of treatment prescribers' patients received a non-acetylated salicylate which can be compared to four percent of comparison prescribers' patients; the one percent difference, although it had the expected sign, was not statistically significant. Thus, there is some evidence of increased utilization of non-acetylated salicylates in one of the two spillover populations, but the evidence is questionable.

**Exhibit 6.2.4**

**HCFA DUR Demonstration Evaluation**
**Impact of RDUR Letters Regarding Misoprostol**
**Percentage of Patients Receiving a Non-Acetylated Salicylate During Follow-Up Period (Drug Started After Letters Mailed)**

| | Tr | Comp | Difference[a] | | |
|---|---|---|---|---|---|
| | pct | pct | pct | sig | 95 pct CI |
| Direct effect: patients named in letters | 1% | 1% | 0% | NS | -3% to 3% |
| Direct effect: patients named in letters who continue with same prescriber | 0% | 1% | -1% | NS | -5% to 2% |
| Spillover 1: other "at risk" patients of prescribers receiving letters | 3% | 1% | 2% | * | 0% to 4% |
| Spillover 2: new patients of prescribers receiving letters | 5% | 4% | 1% | NS | -1% to 4% |

[a] Percentages were rounded for presentation but differences and test statistics were calculated before rounding. The standard test for a difference in proportions was used. * denotes 90 percent confidence that true value of the difference is not zero. ** denotes 95 percent confidence and *** denotes 99 percent confidence. NS indicates a statistically insignificant difference.

*Sources:* Iowa Medicaid claims

The final hypothesis was that the percentage of patients who were at risk for NSAID-induced ulcers and who were potential candidates for misoprostol would fall as a result of the letters. As Exhibit 6.2.5 indicates, 31 percent of patients named in letters remained at risk in the follow-up period as did 29 percent of at-risk patients of comparison prescribers. The difference, which did not have the expected sign, was not statistically significant. If one restricts the sample to patients who continued to see one of their review-period prescribers in the follow-up period, then 37 percent of treatment patients remained at risk compared to 43 percent of comparison patients. This difference did have the expected sign but was again statistically insignificant.

In the first spillover population, at-risk patients of treatment prescribers not named in letters, 25 percent remained at risk in the follow-up period compared to 29 percent of the comparison patients. In the second, new members of the broad risk category, 20 percent of patients of treatment prescribers were at risk by the narrow definition compared to 23 percent of patients of comparison prescribers. Both these differences had the expected sign but were not statistically significant. Thus, one cannot conclude with confidence that the RDUR letters reduced the percentage of patients who were at risk in the follow-up period either as a direct or as a spillover effect.[10]

**Exhibit 6.2.5**

**HCFA DUR Demonstration Evaluation**
**Impact of RDUR Letters Regarding Misoprostol**
**Percentage of Patients At Risk During Follow-Up Period**

| | Tr | Comp | Difference[a] | | |
|---|---|---|---|---|---|
| | pct | pct | pct | sig | 95 pct CI |
| Direct effect: patients named in letters | 31% | 29% | 1% | NS | -10% to 12% |
| Direct effect: patients named in letters who continue with same prescriber | 37% | 43% | -7% | NS | -21% to 8 % |
| Spillover 1: other "at risk" patients of prescribers receiving letters | 25% | 29% | -4% | NS | -21% to 4% |
| Spillover 2: new patients of prescribers receiving letters | 20% | 23% | -2% | NS | - 7 % to 2% |

[a] Percentages were rounded for presentation but differences and test statistics were calculated before rounding. The standard test for a difference in proportions was used. * denotes 90 percent confidence that true value of the difference is not zero. ** denotes 95 percent confidence and *** denotes 99 percent confidence. NS indicates a statistically insignificant difference.

*Sources:* Iowa Medicaid claims

## Conclusion and Discussion

In 1995, the Iowa Medicaid DUR Board mailed letters to 73 prescribers and 83 pharmacies whose patients appeared to be good candidates for misoprostol, the only drug approved by the FDA for the prevention of gastric ulcers caused by NSAIDs. This chapter has explored the impact of those letters on misoprostol utilization and other drug utilization outcomes among patients of the contacted prescribers.

The main finding is that the letters did lead to a statistically significant increase in the use of misoprostol among the patients who were identified in the letters. Specifically, eight percent of patients identified in the letters received misoprostol in the follow-up period while only one percent of comparison patients received it. There was not evidence that the letters induced an increase in percent of patients receiving misoprostol in two potential spillover populations: other patients of treatment prescribers who were at risk during the review period but were not named in the letters and patients of treatment prescribers who were new members of a broadly defined risk group in the follow-up period.

---

[10] It may seem paradoxical that one can conclude that the letters increased the percentage of patients receiving misoprostol but not that they decreased the percentage of patients remaining at risk; this is a function of the large number of outcome categories and the tighter confidence intervals around relatively rare outcomes.

Although the letters and accompanying educational material emphasized prophylaxis via misoprostol, they also mentioned the fact that the use of non-acetylated salicylates was associated with lower ulcer risk than the use of other NSAIDs. This study did not find that the letters led to increased use of non-acetylated salicylates among patient named in letters. It did find some somewhat questionable evidence that they may have increased the use of non-acetylated salicylates among patients in the first spillover population. Finally, there was no indication that the letters led to a decrease in the percentage of patients at risk in the follow-up period among either the population of patients named in letters or either of the spillover populations.

It is important to note that the intervention's effect on misoprostol use, although significant, was small. More than 90 percent of the patients named in the letters did not receive misoprostol in the follow-up period. Within the framework of the current study, there is no way to judge the percentage of at-risk patients for whom misoprostol would have been the optimal therapy. More than half the patients judged at risk of GI bleeds in the review period were no longer at risk in the follow-up period due to other changes in their profiles. Although the Iowa Medicaid DUR Board made every effort to send letters only in cases where the patient appeared to be a good candidate for misoprostol, the facts that the Medicaid record is often incomplete, particularly regarding services not paid by Medicaid, and that a physician's judgment is generally preferred to an absolute standard, no matter how carefully researched, mean that one cannot assume that all the patients who remained at risk according to the criteria should have been taking misoprostol. It *is* safe to assume that the addition of misoprostol did represent an improvement in therapy when it occurred.

This study left some uncertainty regarding spillover effects. It found no evidence of spillover effects on the outcome of primary interest, misoprostol use, and questionable evidence, for one of the two spillover populations, on the secondary outcome, use of non-acetylated salicylates. Ideally, RDUR letters would amplify their benefits by improving therapy for other current and future patients of a physician not only for those patients identified in letters. It should be noted that these interventions were a one-time effort to alter prescribing patterns in these clinical situations. The Iowa DUR Board's RDUR interventions have evolved considerably since 1995; they now repeat letters interventions and/or use supplementary forms of education. It would be worthwhile to determine if letter interventions repeated at regular intervals or a combination of letter interventions with other methods such as targeted education, OPDUR screens, telephone contact, or in-person meetings would change prescribing habits more broadly and lead to spillover effects.

It might also be worthwhile to extend this study, or to design other studies, to examine RDUR letters' effects on the actual clinical outcomes and associated health services utilization that the recommended changes in drug therapy were supposed to improve. For the Iowa misoprostol intervention, the fact that patients named in letters were more likely to receive misoprostol than comparison patients might indicate that they were also less likely to suffer from ulcers and gastropathy later and would have lower associated health services utilization; it would be interesting to know. Moreover, with this information, one could do a complete cost-effectiveness evaluation that took into account the cost of the RDUR program, the cost of all related changes in drug therapy, and the cost of any associated changes in other health services utilization. Unfortunately, these extensions are beyond the scope of the current project.

As it is, this study contributes to a small body of research that demonstrates that Medicaid RDUR letters do make a difference for the subsequent drug therapy of patients named in the letters. It extends that research both by assessing the impact of RDUR directed towards underutilization, as opposed to over utilization, of a target drug and by distinguishing between effects on patients named in letters and on other patients of targeted prescribers. A thorough understanding of the impact of Medicaid RDUR is especially important at this time because many states are currently developing and refining their programs in response to the recent federal mandate.

## 6.3 The Impact of RDUR — Iowa Medicaid DUR Board Letters on Salmeterol

### Introduction

Health problems resulting from less-than-optimal drug therapy are prevalent and severe. Estimates of the percent of new prescriptions containing errors that may result in health problems range from .78 percent to 2.6 percent.[11] A review of hospital admissions suggested that fully 28 percent were caused by drug-related morbidity and mortality; this same study estimated that the cost associated with the management of drug-related morbidity and mortality was $76.6 billion annually [Johnson & Bootman (1995)].

This chapter reports on a quasi-controlled study of an RDUR initiative in Iowa. In the spring of 1995, the Iowa DUR Board sent educational letters to the physicians and pharmacists of Medicaid patients who had claims for salmeterol but not for short-acting bronchodilators (SABDs). (Salmeterol is a long-acting bronchodilator that is not effective in acute asthma attacks; thus, it should always be prescribed in conjunction with a short-acting bronchodilator for use on an as-needed basis during acute episodes.) This study, described in detail below, found that among the patients whose profiles triggered letters, there was an increase in the percentage receiving salmeterol and a short-acting bronchodilator in the follow-up period that can be attributed to the intervention. This study found no evidence that the RDUR letters affected utilization of targeted drugs among new salmeterol patients of the prescribers who received letters.

---

[11] See: Dobie III, Richard L. and Rascati, Karen L. "Documenting the Value of Pharmacist Interventions," *American Pharmacy*, 1994: NS34(5): 50-54; Rupp, M. T., Schondelmeyer, S. W., Wilson G. T., and Krause J. E., "Documenting Prescribing Errors and Pharmacist Interventions in Community Pharmacy Practice," *American Pharmacy*, 1988, NS28(9): 574-580; Rupp, M.T., DeYoung, M., and Schondelmeyer, S. W. , "Prescribing Problems and Pharmacist Interventions in Community Practice," *Medical Care*, 1992, 30(10): 926-9 40.

Note that these estimates only include errors that pharmacists noticed, thus they may understate the actual problem.

### RDUR in Iowa

The Iowa Department of Human Services contracts with the Iowa Foundation for Medical Care to provide Medicaid RDUR; the Iowa Foundation for Medical Care, in turn, subcontracts with the Iowa Pharmacists Association (IPA) for administration of the program. The Iowa Pharmacists Association is unusual in that they develop their own screening routines for their focused studies (described in detail below) rather than contracting with another vendor.[12] Moreover, the Iowa RDUR program is distinguished by ongoing self-evaluation and change with the goal of providing the most effective program possible. This program complies wholeheartedly with the spirit as well as the letter of OBRA 1990.

In 1995, Iowa's Medicaid RDUR program had two components. The first was an ongoing effort to intervene in the "worst of the worst" cases. All patients were assessed for general drug-related risk based on automated utilization screens such as the number of prescriptions, number of prescribers, and number of pharmacies. Patients were then ranked in order of general risk and profiles were screened using a broad range of automated therapeutic criteria. Profiles of patients with the highest risk which also failed at least one therapeutic criteria screen were extracted and subjected to manual review. If the reviewer determined that there might indeed be a problem, the DUR Board sent an educational letter to the prescribers and pharmacists involved in the patient's care for their consideration and comment. The main purpose of these letters was to improve drug therapy and to avoid drug-related problems. An additional benefit of these reviews was that they enabled the Board to evaluate the relevance of proposed criteria in the field and to identify particular prescribing problems that were sufficiently prevalent and severe to merit targeted educational efforts.

The second component of the RDUR program was problem-focused studies, concentrated efforts to address particular prescribing problems. This paper is an evaluation of a focused study conducted in 1995 intended to ensure that patients receiving salmeterol also received a short-acting bronchodilator for use on an as-needed basis in acute asthma attacks. Lags in Abt Associates' access to Medicaid claims data make it necessary to evaluate a 1995 intervention rather than a more recent one. It is important to note that the RDUR intervention under study was one of Iowa's early focused studies and that the program has continued to evolve over the intervening period.

The Iowa Medicaid DUR program continues to perform both "worst-of-the-worst" patient-focused reviews and problem-focused studies. The latter are used to comprehensively address problems that surface during the patient-focused reviews. Some focused studies are repeated in order to provide continuous education and ongoing reinforcement to providers. The Board also analyzes the prescribing patterns of individual physicians relative to the norms for their specialty in order to identify individuals who may benefit from specialized education.

---

[12] Of 48 states for which data is available 30 contract directly with outside vendors for RDUR and 18 use colleges of pharmacy, pharmacists associations, or in-house expertise. It is not clear how many of these agents in the latter group subcontract with outside vendors.

*Salmeterol*

Salmeterol xinafoate (Allen and Hanburys' Serevent®) is a long-acting beta agonist indicated for the maintenance treatment of chronic asthma and the prevention of bronchospasm. Like other bronchodilators, it is available in metered dose inhalers. But it is unlike other bronchodilators in two important regards. First, it is effective for up to 12 hours, a significant convenience, especially for patients with nocturnal symptoms. Second, it does *not* relieve acute asthma symptoms. Physicians prescribing salmeterol should also provide their patients with a short-acting inhaled beta2-agonist (e.g. albuterol) for use on an as-needed basis in acute asthma attacks. Patients must be educated regarding which inhaler to use for maintenance treatment and which for intervention in breathing emergencies. At the same time, patients should be aware that extremely large doses of salmeterol or of both salmeterol and other beta2-agonists carry a risk of adverse cardiovascular effects.

Salmeterol was approved by the FDA on February 4, 1994, and was available to patients beginning in April of that year.

### The Iowa Salmeterol Intervention

In 1994, the Iowa Medicaid DUR Board became concerned that the Iowa Medicaid population's utilization of salmeterol might be less than optimal. These concerns were based on the facts that salmeterol was a new drug with a growing market share, that press reports suggested physicians might not fully understand its mechanism of action, and that the consequences of uncontrolled asthma can be serious. The Summer 1994 edition of *The Bulletin of Medicaid Drug Utilization Review in Iowa* featured a first-page article on salmeterol (Serevent®) reviewing the literature on uses and cautions. Appendix K contains a copy of this article.

Because profile review showed that many Iowa Medicaid salmeterol patients still were not receiving short-acting bronchodilators, the Iowa Medicaid DUR Board decided to launch a focused RDUR initiative on the topic. They identified all patients who, based on Medicaid drug claims, had received salmeterol between October 1, 1994 and March 31, 1995, but had not received a short-acting bronchodilator during this period (see Appendix K for a list of SABDs); 379 patients were flagged by these criteria. The Board then mailed educational letters to the physicians who had prescribed salmeterol for these patients and to pharmacists at the pharmacies that had filled the prescriptions. Copies of these letters are attached as Appendix K. The board also mailed a second letter to the physicians only requesting comments on the named patient and on the intervention. Finally, they re-reviewed the profiles of the identified patients for the six month period June 1 through November 30, 1995, and analyzed patient outcomes and physician comments.[13]

---

[13] Some physicians explained that their patients had no claims for SABDs because their asthma was so well controlled by salmeterol that the SABD although available was so rarely needed that no refill was required. However, research by Allen and Hanburys (maker of Serevent®) indicates that in general salmeterol patients will need an SABD refill within a six month period.

### Research Method

The ultimate goal of RDUR is to improve drug therapy for *patients*. Letters are mailed to physicians and pharmacists because they are in the position to solve patients' drug therapy problems. A patient may be without an SABD either because none was prescribed or because a prescription was written but never filled. In the first instance, a physician can rectify the situation directly; in the second, the physician and pharmacist can stress the importance (and appropriate use) of the "emergency" medication to the patient, thereby influencing behavior.

#### *Research Questions*

This research explored two questions:

1. *Did the RDUR letters provoke changes in therapy for the patients whose profiles triggered them?* Increases in the percentage of these patients who received salmeterol plus an SABD and decreases in the percentage who received salmeterol only would be the most immediate expected effect of the letters.

2. *Were there spillover effects? Were new salmeterol patients of contacted prescribers also more likely to receive SABDs?* Although the letters named individual patients, they might have influenced a physician's prescribing more broadly. Such spillover effects could justify RDUR even if the letters arrived too late to help the patients whose cases initially triggered them. Although the analysis was limited by necessity to spillover effects affecting other Medicaid patients, these benefits could extend to all patients of a contacted prescriber.

#### *Outcomes*

The analysis of Question 1 was structured around four outcomes.

1. *Salmeterol plus SABD* A patient might receive both salmeterol and an SABD following the letter. The letters were specifically intended to increase the probability of this outcome.

2. *Salmeterol only* A patient might continue to receive salmeterol only; the letters were specifically intended to decrease the probability of this outcome.

3. *No salmeterol, other asthma drugs* A patient might stop receiving salmeterol because it was not effective in controlling her asthma. Although the letters were not specifically directed at this outcome, they might decrease its probability if there were some patients who, absent the letters, would have concluded, together with their physicians and pharmacists, that salmeterol was not working for them and stopped it but who, with information from the letters, added an SABD and found the dual therapy worked well.

4. *No salmeterol, no asthma drugs* A patient might also stop salmeterol because she no longer needed asthma therapy.

The analysis of Question 2 was based on patients receiving salmeterol only and so only the first two outcomes were relevant. Notice that a patient must fall into one and only one category. Therefore, an increase in the percentage of patients in one outcome category must be offset by a decrease in the percentage of patients in another.

Patients who were no longer eligible for Medicaid at the start of the follow-up period were by necessity excluded from the analysis. Also excluded were patients who added an SABD in the six weeks immediately following the end of the review period. This outcome could not be attributed to the letters in Iowa because it occurred in the window between the time patients were identified and the time the letters were mailed.

### Analysis Design

Ideally, the study would compare what happened after the letters were sent to what would have happened had they had not been sent. Given that there was no way to be sure what would have happened absent the RDUR letters, the challenge became to construct a reasonable comparison group.

Recall that patients and prescribers were selected for the treatment cohort based on a computerized review of Iowa Medicaid claims between October 1, 1994 and March 31, 1995. This study's approach was to select patients and prescribers for the comparison cohort based on a similar review of Iowa Medicaid claims during an earlier period. The strength of this comparison is that salmeterol prescribers, pharmacists, and patients in Iowa at two points in time should have been very similar to each other; in fact, some of the same individuals may have been involved. The weakness is that there may have been changes over time in information and knowledge regarding salmeterol, particularly as it was a new drug. In order to control for these differences over time, the "differences-in-differences" estimator was used. The estimated RDUR effect was defined as the difference in outcomes between treatment and comparison cohorts in Iowa *minus* any difference in outcomes over the same period in a comparison state. In this way, the estimated effect was cleansed of any change that might be due to broader time trends.

Algebraically, this may be expressed:

$$\text{estimated effect} = (\bar{X}_{IA}^{\text{Cohort 2 (RDUR)}} - \bar{X}_{IA}^{\text{Cohort 1}}) - (\bar{X}_{\text{Comp}}^{\text{Cohort 2}} - \bar{X}_{\text{Comp}}^{\text{Cohort 1}})$$

where $\bar{X}$ is the average of the outcome measure in Iowa or the comparison state for patients in Cohort 2 (the treatment cohort in Iowa) or Cohort 1 (the comparison cohort in Iowa).

### Data Development

This study was based exclusively on drug claims for Medicaid recipients in Iowa, Washington, and Georgia in 1994 and 1995. Data for the study were organized around three six month time periods two months apart:

| | |
|---|---|
| Period 1 | February 1, 1994 - July 31 1994 |
| Period 2 | October 1, 1994 - March 31, 1995 |
| Period 3 | June 1, 1995 - November 30, 1995 |

Periods 2 and 3 were the case-identification and follow-up periods defined by the Iowa Medicaid DUR Board for their intervention. Period 1 was used to identify cases for Iowa comparison cohort and

Period 2 was the follow-up period for these cases. The astute reader may note that Period 1 begins rather early given that salmeterol was only approved in February and only came to market in April. However, this early period was necessary in order to have a complete follow-up period to measure outcomes before the RDUR letters were mailed as the arrival of these letters could contaminate the results.

Thus, the data structure was the following:

| | | |
|---|---|---|
| Cohort 1 | patients identified Period 1 | follow-up Period 2 |
| Cohort 2 | patients identified Period 2 | follow-up Period 3 |

The evaluation had access to data for three potential comparison states: Washington, Georgia, and Maryland. Of these states, Washington is most similar to Iowa in climate and demographics, two factors that affect both the prevalence of asthma in a state and the optimal choice of treatment. However, Washington was also the site of a demonstration of payment for pharmacists' cognitive services during this period and the Peak project, a pharmacy-based initiative to offer intensive and structured care to adolescent and pediatric asthma patients. The presence of these projects could alter outcomes in Washington and make it an unsuitable comparison state. Georgia and Maryland's climate and demographics are relatively different from Iowa, and Maryland also had DUR policies, notably on-line prospective DUR, that distinguished it from Iowa. Georgia did not have any apparently confounding DUR policies.

The evaluation analyzed both the comparison between Iowa and Washington, excluding all patients who went to pharmacies participating in either the cognitive services demonstration or the Peak project, and the comparison between Iowa and Georgia. Summary measures based on the comparison of Iowa with both states were also constructed.

Once the time periods were defined and the comparison states chosen, the first step was to write a simple computer program to flag patients who received salmeterol without a short-acting bronchodilator and to apply it to Iowa Medicaid drug claims during Period 2 in order to test whether it would indeed identify the same patients whose profiles generated letters in Iowa.

The result of this test, shown in Exhibit 6.3.1, was that only eight, (2.3 percent), of the 347 patients identified by the Abt review were not named in Iowa letters. Of the 379 patients whose cases generated letters in Iowa, 40, (11 percent), were not flagged by the Abt claims review. (Abt claims review, which relied on more recent and more complete data than was available in Iowa at the time of the intervention, indicated that the 40 patients actually had received an SABD.) [14]

---

[14] In order to have a clean comparison, all exhibits except Exhibit 8-1 define the Iowa treatment patients/prescribers based on the Abt computerized review and not the Iowa review. The paper's conclusions are not sensitive to this decision.

**Exhibit 6.3.1**

### HCFA DUR Demonstration Evaluation
### Impact of RDUR Letters Regarding Salmeterol
### Comparison of Patients Identified Via Abt Claims Review and Patients Selected for Iowa Intervention

| | Selected Via Abt Claims Review | Not Selected Via Abt Claims Review |
|---|---|---|
| Selected for Iowa Intervention | 339 | 40[a] |
| Not Selected for IA Intervention | 8[b] | |

[a] Abt claims review, based on more recent and complete data, indicated that these 40 patients had received short-acting bronchodilators.

[b] Presumably, these cases were not selected in Iowa because evidence of the salmeterol claim was missing.

*Sources:* The Iowa Medicaid DUR Board (information on the Iowa salmeterol intervention), Medicaid drug claims, 1994-1995.

The next step was to select patients for Cohort 1 in Iowa and Cohorts 1 and 2 in Washington and Georgia by applying the flagging program to the appropriate claims. Then outcomes were categorized for these patients. Also, new salmeterol patients (i.e. patients who received salmeterol in the follow-up but not the review period) were identified and categorized for all flagged prescribers.

#### Utilization of Salmeterol with and without SABDs

This section both gives background on the utilization of salmeterol with and without concurrent utilization of short-acting bronchodilators in the three study states and establishes that their experiences were sufficiently similar to justify using Washington and Georgia as comparison states for Iowa.

In the Iowa Medicaid population, salmeterol utilization grew rapidly during salmeterol's first year on the market and then began to level off. (See Exhibit 6.3.2) The number of salmeterol patients more than doubled between Periods 1 and 2 and increased an additional 11 percent between Periods 2 and 3. By Period 3, there were 804 Medicaid patients filling prescriptions for salmeterol, representing about 2.8 salmeterol patients per thousand Medicaid patients. The growth of the number of salmeterol prescriptions, Medicaid dollars spent on salmeterol, and the number of physicians writing prescriptions for salmeterol roughly paralleled the growth in the number of patients.[15]

In both Washington and Georgia, the number of patients more than doubled between Periods 1 and 2 with the percentage increase in Washington very similar to that in Iowa and the increase in Georgia slightly greater. In both states, the number of salmeterol patients actually fell off slightly between

---

[15] Presumably prescriptions and costs grew faster than the number of patients between Periods 1 and 2 because patients began taking the drug in the middle of Period 1.

Periods 2 and 3 suggesting these states' might have been slightly ahead of Iowa in the product cycle.[16] Growth in the numbers of prescriptions, dollars and prescribers in these two states was consistent with

**Exhibit 6.3.2**

**HCFA DUR Demonstration Evaluation**
**Salmeterol Utilization**

| | Period 1<br>Feb 94-July 94 | Period 2<br>Oct 94-March 95 | | Period 3<br>June 95-Nov 95 | |
|---|---|---|---|---|---|
| | number | number | pct change | number | pct change |
| ***Iowa Medicaid Recipients*** | | | | | |
| patients | 319 | 722 | 126% | 804 | 11% |
| Rxs | 495 | 1650 | 233% | 1900 | 15% |
| dollars | $23,000 | $79,000 | 243% | $96,000 | 22% |
| Medicaid prescribers[a] | 173 | 387 | 124% | 454 | 17% |
| ***Washington Medicaid Recipients*** | | | | | |
| patients | 499 | 1107 | 122% | 973 | -12% |
| Rxs | 742 | 2442 | 229% | 2465 | 1% |
| dollars | $38,000 | $119,000 | 213% | $124,000 | 4% |
| Medicaid prescribers[a] | 233 | 472 | 103% | 468 | -1% |
| ***Georgia Medicaid Recipients*** | | | | | |
| patients | 734 | 1875 | 155% | 1774 | -5% |
| Rxs | 991 | 3730 | 276% | 3673 | -2% |
| dollars | $49,000 | $189,000 | 286% | $196,000 | 4% |
| Medicaid prescribers[a] | 212 | 388 | 83% | 727 | 87% |

a For approximately 10 percent of salmeterol patients in IA and WA and approximately 24 percent of salmeterol patients in GA, the prescriber code is blank or a dummy prescriber code is used. Counts of prescribers only include identifiable prescribers.

*Sources:* Iowa, Washington, and Georgia Medicaid drug claims, 1994-1995 growth in the number of patients.[17] On balance, the general patterns in Washington and Georgia seem sufficiently similar to those in Iowa to justify the comparisons discussed above.

---

[16] It is fairly common for the number of patients using a new drug to rise rapidly, dip slightly, and then level off as physicians, pharmacists, and patients become familiar with the optimal use of the drug.

Despite the drug's labeling, a substantial percentage of salmeterol patients apparently did not have a short-acting bronchodilator available for acute asthma exacerbations. In Iowa, 42 percent of salmeterol patients in Period 1 did not fill a prescription for a short-acting bronchodilator.[18] (See Exhibit 6.3.3.) Surprisingly, this number rose to 47 percent in Period 2 (the review period for the letters) then fell to 41 percent in Period 3, the follow-up period for the letters.

In Washington, the percentage of salmeterol patients not receiving an SABD started at a lower level than in Iowa, 34 percent in Period 1, then rose to 41 percent in Period 2 and fell back to 39 percent in Period 3. In Georgia, this percentage was steady at 47 or 48 percent in all three periods. In all three states, in all three periods, a substantial fraction of salmeterol patients did not receive the SABD despite the clear therapeutic indication, clearly warranting the educational effort launched in Iowa. Moreover, the crude percentages show a relative improvement in Iowa between Periods 2 and 3; this is consistent with the letters having their desired effect.

**Exhibit 6.3.3**

**HCFA DUR Demonstration Evaluation**
**Salmeterol Patients Not Receiving Short-Acting Bronchodilators**

| | Period 1<br>Feb 94-July 94 | Period 2<br>Oct 94-March 95 | Period 3<br>June 95-Nov 95 |
|---|---|---|---|
| *Iowa Medicaid Patients* | | | |
| Salmeterol patients | 319 | 722 | 804 |
| Without SABD (number) | 134 | 339 | 327 |
| Without SABD (percent) | 42% | 47% | 41% |
| *Washington Medicaid Patients* | | | |
| Salmeterol patients | 499 | 1107 | 973 |
| Without SABD (number) | 169 | 453 | 380 |
| Without SABD (percent) | 34% | 41% | 39% |
| *Georgia Medicaid Patients* | | | |
| Salmeterol patients | 734 | 1875 | 1774 |
| Without SABD (number) | 345 | 875 | 850 |
| Without SABD (percent) | 47% | 47% | 48% |

*Sources:* Iowa, Washington, and Georgia Medicaid drug claims, 1994-1995

---

[17] The large amount of missing data in Georgia may explain the increase in the number of identifiable prescribers between Periods 2 and 3.

[18] Statistics based on this data will overstate or understate the problem to the extent that patients pay for their SABDs or salmeterol out of pocket or via other insurance. Although noteworthy, these problems should not undermine the general conclusions of the research. (They do not bias the difference-in-difference estimators)

## Effects of the Iowa Intervention

### Patients Whose Cases Generated Letters

The direct effect of the RDUR letters should be to improve drug therapy for the patients whose profiles generated the letters. In order to estimate this effect, outcomes for treatment patients, i.e. patients whose profiles generated letters, were compared to outcomes for the comparison group, patients who received salmeterol without a short-acting bronchodilator in Iowa during Period 1 and in the two comparison states during Periods 1 and 2.

In Iowa, only three percent of the patients flagged for salmeterol without an SABD during Period 1 received salmeterol plus an SABD in Period 2. (See Exhibit 6.3.4.) More than a third of these patients, 39 percent, were still receiving salmeterol only. Fifty-eight percent had stopped salmeterol altogether; approximately half of these were taking other Medicaid-paid asthma drugs, approximately half were not.

On the one hand, the problem of salmeterol without an SABD was not self-correcting; in the absence of prompting, the vast majority of patients did not start the dual therapy. On the other, the problem was not static; more than half of the patients stopped salmeterol.

Twelve percent of patients flagged in Period 2 (whose physicians and pharmacists received letters) received the combination therapy (salmeterol plus an SABD) in Period 3. Thirty-four percent received salmeterol only and 58 percent had stopped the drug. Thus, there was an increase of nine points in the percentage of patients receiving the combination therapy offset not by a decline in the percentage of patients who continued to receive salmeterol only but rather by a decline in the percentage of patients who stopped taking salmeterol.

One may wonder whether this small change was indeed the result of the RDUR letters or part of a general improvement over time in physicians', pharmacists', and patients' understanding of the new drug. The next two sections of the exhibit show the corresponding percentages for the states of Washington and Georgia during the same two time periods.

Five percent of patients flagged in Washington for Cohort 1 were taking the combination therapy in the follow-up period and 25 percent were still receiving salmeterol only.[19] Eight percent of Washington patients flagged for Cohort 2 received both salmeterol and an SABD in Period 3 while 24 percent continued to receive salmeterol only.

In Georgia, nine percent of patients flagged for Cohort 1 received the combination therapy in the follow-up period while 32 percent remained at risk. In this state, there was very little change between the first and second cohorts. In the second cohort, eight percent of patients received the dual therapy in the follow-up period and 32 percent remained at risk.

---

[19] In this section, the phrase "flagged patients" will be used to refer to patients who were selected by the flagging program as receiving salmeterol without a short-acting bronchodilator. The phrase "flagged prescribers" will be used to refer to the physicians who prescribed salmeterol for these patients. This shorthand greatly simplifies the discussion.

**Exhibit 6.3.4**

**HCFA DUR Demonstration Evaluation**
**Outcomes in Follow-Up Period for Patients who Received Salmeterol without a Short-Acting Bronchodilator in Review Period**

| | Cohort 1 Cases Defined Period 1, Follow-Up Period 2 | | Cohort 2 Cases Defined Period 2, Follow-Up Period 3 | |
|---|---|---|---|---|
| | number | pct | number | pct |
| ***Iowa Medicaid Patients*** | | | | |
| SABD > 6 wks | 3 | 3% | 27 | 12% |
| salmeterol only | 36 | 39% | 94 | 41% |
| other asthma drug | 26 | 28% | 45 | 20% |
| no asthma drug | 28 | 30% | 64 | 28% |
| total | 93 | 100% | 230 | 100% |
| ***Washington Medicaid Patients*** | | | | |
| SABD > 6 wks | 6 | 5% | 23 | 8% |
| salmeterol only | 30 | 25% | 97 | 34% |
| other asthma drug | 34 | 29% | 51 | 18% |
| no asthma drug | 48 | 41% | 116 | 40% |
| total | 118 | 100% | 287 | 100% |
| ***Georgia Medicaid Patients*** | | | | |
| SABD > 6 wks | 14 | 9% | 31 | 8% |
| salmeterol only | 51 | 32% | 119 | 32% |
| other asthma drug | 48 | 30% | 93 | 25% |
| no asthma drug | 45 | 28% | 125 | 34% |
| total | 158 | 100% | 368 | 100% |

***Differences in Differences Estimate***

| | estimate (percentage points) | 95 pct conf interval (percentage points) |
|---|---|---|
| ***SABD > 6 weeks*** | | |
| Iowa vs. Washington | 6 | (-2, 13) |
| Iowa vs. Georgia | 9 | (1, 16) |
| Iowa vs. Washington and Georgia | 7 | (1, 14) |
| ***salmeterol only*** | | |
| Iowa vs. Washington | -6 | (-21, 9) |
| Iowa vs. Georgia | 2 | (-13, 17) |
| Iowa vs. Washington and Georgia | -1 | (-15, 12) |

*Note:* Exhibit excluded patients whose salmeterol prescriber could not be identified (approximately 10 percent of patients in IA and WA and 24 percent of patients in GA) and patients who began using an SABD within six weeks of the end of the review period, i.e. before the letters were mailed in Iowa (less than 10 percent of patients in all states and time periods.)

*Sources:* Iowa, Washington, and Georgia Medicaid drug claims, 1994-1995

The bottom section of the table reports the difference-in-difference estimates of the intervention effects and a 95 percent confidence interval. The letters were expected to increase the percentage of patients who received the combination therapy. Based on the comparison of Iowa and Washington, the intervention led to a six percentage point increase in the percentage of patients receiving the dual therapy. In Iowa, the percentage of patients in this category had increased by nine points while in Washington, it had increased three points. Moreover, the absolute percentage of patients receiving the combination therapy was higher in Iowa than in Washington after the letters were mailed; this reinforces the conclusion that the difference-in-difference estimate represents a real effect and not a natural catching-up process. The 95 percent confidence interval for this estimate is -2 to 13 percentage points and includes zero; thus, despite the positive estimate, one cannot reject the null hypothesis that the letters had no effect on this outcome.[20]

Based on the comparison of Iowa and Georgia, the intervention led to a nine percentage point increase in the percentage of patients receiving the dual therapy. Here the nine point increase in Iowa was compared to a one point decrease in Georgia. Again, in absolute terms, the percentage of patients receiving the dual therapy was higher in Iowa than in Georgia in the follow-up period. In this case, the 95 percent confidence interval for this estimate was one to sixteen percentage points and did not include zero, enabling one to reject the null hypothesis of no effect.

A summary measure can be obtained via the comparison of Iowa with Washington and Georgia combined. This estimate is seven percent with a confidence interval of one to fourteen percentage points. Although the confidence interval is wide making a precise estimate impossible, one can conclude with confidence that the letters had a positive effect on the percentage of patients who received salmeterol in conjunction with an SABD.

The letters were also expected to have a negative effect on the percentage of patients who remained at risk, that is who received salmeterol without the rescue medication. Based on the comparison of Iowa and Washington, the intervention led to a six point decrease in the percentage of patients remaining at risk. However, this estimate is driven by an unexpected eleven point increase in Washington counterbalanced by a two point increase in Iowa, not by the expected decrease in Iowa. In addition, the percentage of patients at risk in the follow-up period in Iowa is still higher than the corresponding percentage in Washington. The confidence interval for this estimate includes zero with a wide margin on both sides.

The estimate of the intervention effect based on the comparison of Iowa and Georgia is two points due to the two point increase in Iowa and the absence of change in Georgia. Again, the percentage of patients remaining at risk in Iowa in the intervention period surpasses that in Georgia and the wide confidence interval includes zero. The summary measure too supports the conclusion that one cannot reject the null hypothesis of no effect.

---

[20] Because the sample included all Medicaid patients in Iowa and Washington, the means were exact measures (not estimates) for those two states. The exact measures from Iowa can be considered an estimate of what would have happened if the RDUR policy had been nationwide and the exact measure from Washington can be considered an estimate of what would have happened absent RDUR. In that sense the measures were estimates and hypothesis testing was appropriate.

It may at first seem paradoxical to conclude that the letters raised the percentage of patients receiving both salmeterol and an SABD but did not lower the percentage of patients remaining at risk. The resolution lies in the fact that the letters apparently decreased the percentage of patients stopping salmeterol. One can speculate that, in the absence of the letters, some patients, together with their physicians and pharmacists, would have concluded that salmeterol was not the right drug for them because it was not effective in an acute asthma attack and discontinued its use. However, with the information in the letters, these patients and their treatment teams would realize that salmeterol was indeed effective when used in combination with an SABD.

If anything, it is surprising that the estimated changes in SABD use were so small. Often RDUR seeks to curb over utilization, to promote the substitution of less costly, therapeutically equivalent drugs, or to advocate a drug with unpopular side effects (misoprostol). In these cases, some patient or prescriber resistance to the course of action suggested is understandable. In the case of salmeterol, it is difficult to fathom why patients, prescribers, and pharmacists would not choose to add the SABD given the information in the letter. One has to wonder either whether the Medicaid drug record was incomplete and patients actually did receive the dual therapy or whether the letters went astray or were ignored so that, in fact, in some cases, the information was not conveyed. It is safe to assume that the change to dual therapy represented an improvement when it occurred.

## Effects on Other Patients

The estimates presented in the previous section suggested that the RDUR letters had a positive impact on the patients whose profiles provoked the letters. In principle, the RDUR letters might have affected not only those patients but also other patients of the contacted prescriber.[21] Exhibit 6.3.5 presents outcomes for new salmeterol patients of flagged prescribers. In Iowa, in Period 2, 43 percent of these patients received salmeterol without a short-acting bronchodilator. In Period 3, when the prescribers would have received letters, this figure had fallen to 37 percent. However, there was a corresponding change in Washington. In Period 2, 49 percent of new patients of flagged prescribers did not receive an SABD while in Period 3, this was the case for 35 percent of patients. In Georgia, the percentage of flagged prescribers' new salmeterol patients who did not receive an SABD remained roughly constant, 46 percent in Period 2, 47 percent in Period 3.

Thus, based on the Iowa/Washington comparison, the differences-in-differences estimator of the treatment effect was an eight point *decrease* in the percentage of patients receiving an SABD while the estimator based on the Iowa/Georgia comparison was a six point increase. The estimator based on comparing Iowa to both states was two percentage points. All estimators have wide confidence intervals that include zero; thus, one cannot reject the null hypothesis that the letters had no effect on these spillover populations.

---

[21] Spillover effects may work through pharmacists as well, but are not explored in this paper. To some extent, the issue of spillover effects is intertwined with the issue of persistence. If there are spillover effects and if physicians who exhibit a questionable prescribing pattern regarding one patient are also especially likely to exhibit it with other patients, then RDUR may be an effective way to target these providers and alter their behavior. Appendix D analyzes whether there was persistence in the prescribing pattern of salmeterol without an SABD.

**Exhibit 6.3.5**

**HCFA DUR Demonstration Evaluation**
**Outcomes in Follow-Up Period for New Salmeterol Patients of Flagged Prescribers**

| | Pbrs Identified Period 1, Follow-Up Period 2 | | Pbrs Identified Period 2, Follow-Up Period 3 | |
|---|---|---|---|---|
| | number | pct | number | pct |
| ***Iowa Medicaid Patients*** | | | | |
| SABD > 6 wks | 49 | 57% | 91 | 63% |
| salmeterol only | 37 | 43% | 53 | 37% |
| total | 86 | | 144 | |
| ***Washington Medicaid Patients*** | | | | |
| SABD > 6 wks | 69 | 51% | 77 | 65% |
| salmeterol only | 65 | 49% | 41 | 35% |
| total | 134 | | 118 | |
| ***Georgia Medicaid Patients*** | | | | |
| SABD > 6 wks | 129 | 46% | 138 | 47% |
| salmeterol only | 150 | 54% | 156 | 53% |
| total | 279 | | 294 | |

***Differences in Differences Estimate***

| | estimate (percentage points) | 95 pct conf interval (percentage points) |
|---|---|---|
| ***SABD > 6 weeks*** | | |
| Iowa vs. Washington | -8 | (-25,10) |
| Iowa vs. Georgia | 6 | (-10, 21) |
| Iowa vs. Washington and Georgia | 2 | (-13, 17) |

*Sources:* Iowa and Washington Medicaid drug claims, 1994-1995

## Conclusion

In the spring of 1995, the Iowa DUR Board contacted prescribers and pharmacists of patients who had Medicaid-paid drug claims for salmeterol but had no claims for a short-acting bronchodilator. These letters were intended to be educational and explained that salmeterol was a long-acting bronchodilator that was of no use in acute asthma attacks; consequently, asthma patients receiving salmeterol should also have a short-acting bronchodilator available for use on an as-needed basis during acute attacks.

This paper is an in-depth study of that intervention intended less to "grade" the Iowa intervention and more to develop lessons for future RDUR in all states. Results of this analysis indicate that, among patients whose cases generated letters, the RDUR letters caused an increase in the percentage of patients receiving salmeterol plus a short-acting bronchodilator in the follow-up period. This paper finds no evidence that physicians who received an RDUR letter about salmeterol prescribing for one patient were more likely to alter their salmeterol prescribing, i.e. add an SABD, for their other, new salmeterol patients.

It is encouraging that this research indicated that RDUR can have a positive effect on patient outcomes; it joins several recent studies in that conclusion. It is, however, discouraging that this research did not demonstrate an impact that extends beyond the individual patient named in the letter. In general, RDUR programs have the dual goals of improving therapy for individual patients and changing providers' behavior more generally. The Iowa Medicaid DUR Board now does repeated interventions on issues of importance. Perhaps, over a longer horizon, these interventions will be able to change prescribing behavior more broadly.

## 6.4 The Impact of RDUR — Maryland RDUR Letters on Zantac Use

### Introduction

The objective of this analysis is to assess the effectiveness of a Medicaid RDUR letter intervention designed to modify prescribing of H2 Blockers. The letter intervention, implemented in September, 1994, was intended to reduce long-term use of H2 Blockers at full dosage levels in Maryland. A quasi-experimental study design is used with a comparison group consisting of a set of patients whose prescribers had been targeted to receive the letter intervention, but had not yet received the letters .

The results indicate that three months after the letter intervention, treatment group patients were significantly less likely than comparison group patients to refill prescriptions for the H2 Blocker Zantac (rantidine). Although this effect was quite large in magnitude, data and other limitations precluded a thorough analysis of confounding factors, such as a concurrent academic detailing program. Nevertheless, the results of this study are consistent with the results in the published literature; Medicaid DUR letter interventions are an effective way of modifying prescribing of H2 Blockers.

## Drug Utilization Review in Maryland

Maryland initiated Medicaid drug utilization review (DUR) in 1986, four years before Congress mandated retrospective and prospective drug utilization for all state Medicaid programs. In the last ten years, Maryland DUR has evolved from a retrospective system based on letter interventions to a diverse program involving academic detailing, on-line prospective drug utilization review (OPDUR), and on-going refinement of clinical criteria. Despite the changes in the program, Maryland DUR has focused on educational interventions throughout its history.

### Retrospective Drug Utilization Review

The original Maryland drug utilization review program was designed to retrospectively screen paid Medicaid claims for therapeutic problems. To implement the program, the Maryland Department of Health and Mental Hygiene established contracts with two organizations — the Maryland Pharmacists Association (MPA) and Health Information Designs (HID). Health Information Designs was responsible for the computer generated screening of patients' medical and pharmacy profiles. Patient profiles that failed to meet computer screening criteria then were screened manually by physicians and pharmacists at the Maryland Pharmacists Association [Baylis (1994)].

The rationale for the dual computer and manual screening process was that clinical expertise was needed both to verify the potential therapeutic problem generated by the computer program and to identify new problems that the computer program might have missed. Bayliss (1994) reports that rates of computer-generated false positives were quite high, ranging from 65-70% of the approximately 2000 profiles manually reviewed in a single month [Bayliss (1994)]. State DUR reports, however, suggest that these false positive rates typically were even higher and that the rates varied by drug problem. In 1993-1994, for example, only 9.5% of the 5197 computer generated profiles for long-term use of H2 Blockers led to intervention letters after manual review. Letters designed to change prescribing of H2 Blockers, however, still comprised 35.8% of total letters sent in 1993-1994.

Patient profiles that failed both computer based and manual screening resulted in educational intervention letters sent to prescribers and pharmacists. The tone of the educational letters was informative rather than punitive. The letters included a description of the therapeutic problem and, importantly, the names of the patient(s) involved. For example, in the case of long term/high dose use of H2 Blockers, the intervention letter advised the patient's physician and pharmacist that a particular patient had been taking an H2 Blocker at full therapeutic dose for a number of months. Physicians and pharmacists were encouraged to evaluate the patient's record and to reply to DUR committee with comments or concerns.

The Maryland retrospective drug utilization review letter intervention program targeted a number of drug therapy problems including drug/drug interactions, over and under-utilization, drug/disease interactions, and drug duplication. Unfortunately, information is not available regarding "hit rates' or the frequency of these drug problems in the Medicaid population [Maryland DUR Annual Report (1993)]. Consequently, with the information available from the original Maryland DUR program, one cannot come to any conclusions about the magnitude of the drug problems addressed.

Additionally, because Maryland's DUR program was enacted before OBRA 1990, Maryland's DUR Board was organized after the RDUR program already had been operating for six years. The DUR Board therefore was not involved in the initial program design and criteria development [Maryland DUR Annual Report (1993)]. The H2 Blocker criteria that the Maryland Pharmacists Association and Health Information Designs had been using eventually was evaluated and approved by the DUR committee in January 1993 [Maryland DUR Screening Criteria (1995-1996)]. Even so, information is quite limited on the criteria development and implementation process used in the RDUR program before the DUR Board was created. Perhaps for this reason, Maryland's RDUR letter intervention program was never evaluated prior its termination in 1995.

### Prospective Drug Utilization Review

Maryland also implemented an on-line prospective drug utilization review (OPDUR) system in 1993. In contrast to the RDUR system described previously, Maryland's DUR board was involved in the planning and implementation of the OPDUR program. The DUR board continues to provide oversight and analytical support to the OPDUR program. Because Maryland terminated its RDUR letter intervention program in mid 1995, the OPDUR system has become the focus of the state's drug utilization review efforts. OPDUR does not include a screen for long-term use of H2 Blockers, making it unlikely that OPDUR will confound the effects of RDUR in this analysis [personal communication with Ilene Zuckerman, 10-30-97].

### Other Initiatives Designed to Improve Prescribing

Academic detailing and informational newsletters also have been used to improve prescribing practices in Maryland. As part of the Maryland Pharmacist Association's RDUR effort, trained pharmacists offered face-to-face and telephone interventions to targeted prescribers. In September 1994, 22 face-to-face academic detailing sessions and four telephone/mail interventions were conducted with physicians identified for H2 Blocker prescribing issues. It is possible that these physicians also received the RDUR educational letters that are the focus of this analysis. Unfortunately, information is not available regarding when all of the academic detailing sessions took place or which prescribers received both academic detailing and RDUR letters. The potential consequences of this issue are discussed in a later section.

In 1995, Maryland initiated a more formal academic detailing program called the Pharmaceutical Care Services Program. As part of the program, clinical pharmacists rotated through primary care Medicaid clinics. The pharmacists led in-service training sessions, educated patients, and participated in prescribing decisions. Although this program was in place during the time period of interest in this analysis, available records do not show any evidence that long-term use of H2 Blockers was the focus of an in-service training [Maryland DUR Annual Report (1995)].

All Medicaid providers received annual educational newsletters in 1994 and 1995. The topics of the newsletters in these two years were drug use review and anti-hypertensive drug therapy. In 1992, an annual newsletter focussed on cost-effective prescribing of H2 Blockers. The mailing of this newsletter therefore predates the time periods of interest in this analysis.

#### *The H2 Blocker Long-Term/High Dose Letter Intervention*

Zantac (rantidine) is one of four commercially available H2 Blockers used to prevent and treat ulcers [Mullen (1990)]. As a class, H2 Blockers are considered very safe and effective in relieving ulcer pain and promoting ulcer healing in the short term. Generally, a daily 300mg dose of Zantac taken for four to eight weeks is expected to heal 95% of duodenal ulcers. Similarly, the other H2 Blockers (Tagamet, Axid and Pepcid) are normally prescribed for a duration of 4-8 weeks [Rakel (1996)].

Pharmaceutical companies have been heavily marketing Zantac both to providers and directly to patients. Nevertheless, for patients with benign, active ulcers, Zantac use for more than six weeks generally is not indicated. Furthermore, the safety of Zantac used for more than eight weeks has not been established for patients with benign, uncomplicated ulcerative disease. Long-term use of Zantac and other H2 Blockers may not be appropriate therapy even for patients with more persistent ulcers. While some elderly and severely ill patients can benefit from long-term, maintenance therapy at a lower dosage level, in many cases intermittent drug treatment is preferable to maintenance therapy for recurrent ulcers [Rakel (1996)].

In addition to its questionable clinical appropriateness, long-term use of H2 Blockers can be very costly for Medicaid. In 1990, Maryland Medicaid spent $6 million on H2 Blockers alone. [Equal Time (1992)] Of the H2 Blockers, Zantac is the most expensive. Exhibit 6.4.1 below summarizes the costs of a 30-day supply of H2 Blockers to Maryland Medicaid in 1992.

**Exhibit 6.4.1**

**HCFA DUR Demonstration Evaluation**
**Costs of 30 Day Supply of H2 Blockers to Maryland Medicaid, 1992**

| | Acute Dose | Maintenance Dose |
|---|---|---|
| Tagamet (Cimetidine) | $68.24 | $34.12 |
| Pepcid (Famotidine) | $71.28 | $35.64 |
| Axid (Nizatidine) | $73.11 | $36.55 |
| Zantac (Ranitidine) | $81.07 | $40.53 |

a Acute dose for Zantac and Axid is 150 mg bid or 300 mg hs. Acute dose for Pepcid is 20 mg bid or 40 mg hs. Acute dose for Tagement is 300 mg qid, 400 mg bid, or 800 mg hs. Maintenance dose for Zantac and Axid is 150 mg hs. Maintenance dose for Tagamet is 400 mg hs. Maintenance dose for Pepcid is 20 mg hs. All prices based on wholesale aquisition cost.

*Sources:* Equal Time: Prescribing with the Whole Story, Vol. 1 No. 1 April 1992

The preceding discussion suggests that long-term use of H2 Blockers, particularly at acute dosage levels, can be an expensive as well as a clinically inappropriate treatment strategy in cases of benign, uncomplicated ulcer. The Maryland RDUR program addressed this issue by sending intervention letters to the providers of Medicaid patients who had been using H2 Blockers at acute dosage levels for more than 56 days. This analysis is limited to patients who were targeted for use of Zantac at acute dosage levels for more than 56 days.

The H2 Blocker Long-Term/High Dose letter intervention was one of the Maryland RDUR program's largest and most frequent letter interventions. Between 1993-1995, the RDUR program sent 661 intervention letters related to long-term use of H2 Blockers [Annual Reports (1994, 1995)]. The majority of the letters were sent for long-term use of Zantac.

### Study Design: Direct Impact of the RDUR Letter Intervention

One of the goals of this analysis is to evaluate the effectiveness of the H2 Blocker intervention letters in changing Zantac prescribing decisions *for patients who were named in the RDUR letters*. These effects are considered to be the direct impact of the letters. Although the interventions letter did not include specific therapeutic suggestions, it is likely that prescribers responded to the letters by discontinuing Zantac use, by reducing the Zantac dosage level, or by switching the patient to another H2 Blocker. It is also likely that some prescribers decided that for his/her patient(s), long-term use of Zantac at an acute dosage level was appropriate therapy.

A simple analysis of Zantac use before and after the letter intervention may not capture the true direct impact of the intervention. Over time, it is likely that many prescribers would discontinue use of Zantac or reduce Zantac dosage independently. Furthermore, other educational interventions, policies, or practice patterns that change over time might affect Zantac prescriptions. A before/after analysis would not be able to separate the impact of the letter intervention from these confounding factors.

To better estimate the true direct effect of the intervention letters on prescribing practices, a quasi-experimental study design is used. The design compares a treatment group's Zantac prescriptions after the letter intervention to the Zantac prescriptions in a comparison group that was targeted for a letter intervention but had not yet received it. Medicaid patients identified as long-term Zantac users are divided into two groups (treatment and comparison) based on when they were identified by the Maryland Pharmacists' Association and when their providers were sent letters. The treatment group was identified in 7/94 and sent letters between 9/8/94 and 9/15/94. The comparison group was identified in 1/95 sent letters between 3/7/95 and 3/9/95. Outcomes for the treatment and comparison groups are compared during a follow-up period that is defined in the following paragraphs.

The treatment group in this analysis is a group of 81 Medicaid patients whose records were manually reviewed and selected for intervention by the Maryland Pharmacists Association (MPA) in July, 1994. According to the Maryland Pharmacists Association's records, treatment group patients had been using Zantac at acute dosage levels for more than 56 days prior to the review process.[22] The treatment group's providers received educational letters in early September, 1994, approximately two months after MPA's manual review.

The comparison group consists of 96 long-term Zantac users who were identified by MPA in January, 1995. The comparison group was identified using the same criteria that was used to identify the treatment group — use of Zantac for more than 56 days.[23] Letters were sent to the comparison group's providers in early March, 1995, approximately two months after MPA's manual review. The treatment and comparison groups are mutually exclusive.

The analysis compares treatment and comparison group patients' frequency of Zantac prescriptions, dosage of Zantac prescriptions and prescriptions for other H2 Blockers during the follow-up period defined below:

- For the treatment group, the follow-up period is 10/1/94 - 1/1/95. This is the three-month period after the letters were sent.[24] (The letters were sent 9/8/94 - 9/15/94)
- For the comparison group, the follow-up period is 1/1/95 - 4/1/95. This is the three-month period after the patients were targeted for letters but before the letters were received.[25] (The letters were sent 3/7/95 - 3/9/95)

---

[22] To ensure comparability with the comparison group, the treatment group was limited to patients who had been using Zantac for 4-6 months when they were targeted by MPA in July, 1994. This way, at the beginning of the follow-up period in October 1994, the treatment group presumably had been using Zantac for about 6-8 months.

[23] To ensure comparability with the treatment group, the comparison group was limited to patients who had been using Zantac for 6-8 months when they were identified in January 1995. For the comparison group, the follow-up period begins in January, 1995.

[24] It is assumed that prescribers would not be able to respond to the letters until October 1 at the earliest.

[25] Even though the comparison group's letters were sent at the beginning of March, it is assumed that prescribers were not able to receive, read, and react to the letters until April 1 at the earliest. This assumption allowed the follow-up period to be extended to a full three months.

|______________________|___follow-up period____| Treatment
7/94 (prescribers selected) 9/94 (letters sent) 1/95
|______________________|______________________|__follow-up period_______| Comparison
1/95 (prescribers selected) 3/95 (letters sent)

During the follow-up period, simple treatment/comparison group comparisons and probit models are used to address the following three research questions:

- In each of the three months of the follow-up period, were patients named in the intervention letters less likely to refill Zantac prescriptions relative to the comparison group?
- In each of the three months of the follow-up period, were patients named in the intervention letters more likely than comparison group patients to refill Zantac prescriptions at a lower dosage level?
- In each of the three months of the follow-up period, were patients named in the intervention letters more likely to refill prescriptions for Zantac substitutes (Tagamet (cimetidine), Pepcid (famotidine), or Axid (nizatidine)) relative to the comparison group?

### Study Design: Spillover Impact of the RDUR Letter Intervention

Another objective of this study is to assess the effectiveness of the Zantac RDUR letter interventions in modifying prescriptions of "spillover populations", defined as patients of targeted prescribers who were not named in the intervention letters but who were at some point considered to be at risk. The focus on both direct and spillover effects is needed to assess whether or not RDUR letter interventions lead to lasting, broad prescribing improvements that benefit at-risk patients beyond those patients who are named in the letters.

To assess the spillover impact of the RDUR letters, Medicaid claims were used to identify patients who were long-term Zantac users, who had a physician who had received the letter intervention, but who had not been named in the letter intervention. Both the spill-over group and the comparison group were limited to patients who had been using Zantac for two consecutive months in the three months before the intervention (in the case of the spillover group), or before being targeted for an intervention (in the case of the comparison group).

Because only limited information is available about the screening process that was used to select physicians for the letter intervention, it is not clear why some at-risk patients were not named in the letters their targeted physicians received. This lack of information is a limitation in the analysis of spill-over effects; it is possible that spill-over patients (at-risk patients of targeted physicians who were not named in letters) differ from treatment group patients (at-risk patients of targeted physicians who were named in letters) in observable and unobservable ways that affect whether or not they continue to use Zantac.[26]

### Descriptive Statistics

Exhibit 6.4.2 below compares basic characteristics of the treatment group (N = 81) and the comparison group (N = 96).

**Exhibit 6.3.2**

**HCFA DUR Demonstration Evaluation**
**Treatment and Comparison Group Basic Characteristics**

| | Treatment Group | Comparison Group |
|---|---|---|
| Average Age*** | 77.1 | 57.3 |
| % Female*** | 83% | 59% |
| % Caucasian*** | 64% | 67% |
| Average Number of Months Zantac Use During Baseline Period *** | 2.01 | 2.54 |

*** indicates statistically significant differences at the .01 level, two sample t-test and proportions tests used to test differences between groups

*Sources:* Mid-Atlantic DUR, Maryland Medicaid claims

As Exhibit 6.3.2 indicates, the differences between the treatment and comparison groups are quite striking. The treatment group consists of more elderly and female patients compared to the comparison group. Based on available information about Maryland RDUR, it does not appear that the H2 Blocker letter interventions were designed to target specific age groups or a specific gender. It is quite surprising, therefore, that the demographic characteristics of the two groups are so different.

The clinical literature suggests that the elderly may be more likely than others to benefit from long-term H2 Blocker therapy. Because the treatment group patients are older than the comparison group patients, a comparison between treatment and comparison outcomes may understate the true impact of the letters. For example, in the comparison group, many prescribers may have decided to discontinue use of Zantac during the follow-up months for younger patients. In the treatment group, prescribers may have decided that despite the letter intervention, long-term Zantac therapy made sense for their elderly patients. To account for this problem, age will be included as a covariate in the probit models described in the next section.

---

[26] A comparison of the treatment group and the spill-over group populations revealed that, compared to the treatment group, the spill-over population was somewhat younger, less likely to be female, more likely to be Caucasian, and were using Zantac for a slightly longer period of time.

### Results: Direct Impact of the RDUR Letter Intervention

Exhibit 6.4.3 below presents trends in Zantac refill prescriptions for the treatment and comparison groups during the three follow-up months. Zantac refills in the comparison group remain quite stable over the three month follow-up period. In the treatment group, however, Zantac refill prescriptions drop from 62% in the first follow-up period to 44% by the third follow-up period. This large decrease suggests that the letter intervention may have encouraged prescribers to discontinue their patients' Zantac prescriptions.

**Exhibit 6.4.3**

**HCFA DUR Demonstration Evaluation**
**Zantac Refill Prescriptions in Follow-Up Months**

| | Follow-Up Month 1: Number and % Patients With At Least One Zantac Refill | Follow-Up Month 2: Number and % Patients With At Least One Zantac Refill | Follow-Up Month 3: Number and % Patients With At Least One Zantac Refill |
|---|---|---|---|
| Treatment Group | 52<br>62% | 50<br>60% | 37<br>44% |
| Comparison Group | 73<br>75% | 74<br>76% | 72<br>73% |

The treatment group received the letter intervention in 9/94. The follow-up period for this group is 10/94-12/94. The comparison group's follow-up period is 1/95-3/95. Patients may have refilled more than one Zantac prescription in a single month.

*Sources:* Maryland Medicaid, Mid-Atlantic DUR

The intervention letters also may have led to decreases in therapeutic dosage levels. The text of the letter includes a statement indicating that it may be appropriate to lower the Zantac dosage to half of the current dosage. Interestingly, as displayed in Exhibit 6.4.4 below, average dosage levels actually increased in the treatment group during the follow-up period. It is possible that the letter intervention encouraged prescribers to remove their healthiest patients from Zantac entirely, leaving patients with severe ulcer disease and high dosage levels in the treatment group by the third month. This scenario is consistent with the clinical literature which suggests that long-term H2 Blocker use may be appropriate for patients with complicated and persistent ulcers.

**Exhibit 6.4.4**

**HCFA DUR DemonstrationEvaluation**
**Dosage of Zantac Refill Prescriptions in Follow-Up Months**

| | Follow-Up Month 1: Number and Average Dosage Level of Zantac Users | Follow-Up Month 2: Number and Average Dosage Level of Zantac Users | Follow-Up Month 3: Number and Average Dosage Level of Zantac Users |
|---|---|---|---|
| Treatment Group | 52<br>1.8 | 50<br>1.7 | 37<br>2.2 |
| Comparison Group | 72<br>2.3 | 73<br>2.3 | 70<br>2.0 |

The treatment group received the letter intervention in 9/94. The follow-up period for this group is 10/94-12/94. The comparison group's follow-up period is 1/95-3/95. Patients may have refilled more than one Zantac prescription in a single month. Dosage level defined as pills per day (total pills/days supply). Patients with dosage levels above 20 pills per day were dropped from the analysis because of potential coding errors.

*Sources:* Maryland Medicaid, Mid-Atlantic DUR

Prescribers also may have reacted to the letter intervention by changing their patients' prescriptions from Zantac to another H2 Blocker. Although Zantac's cost was not mentioned in the letter, many prescribers might be aware that less costly alternatives to Zantac exist. These prescribers may have inferred that the objective of the letter was to lower Medicaid costs. Exhibit 6.4.5 below displays the numbers of treatment and comparison patients who filled prescriptions for Zantac substitutes in the three follow-up months. In both groups, the numbers are very small, suggesting that few prescribers responded to the intervention by prescribing Zantac substitutes.

**Exhibit 6.4.5**

**HCFA DUR Demonstration Evaluation**
**Zantac Substitute (Tagamet, Pepcid, or Axid) Prescriptions in Follow-Up Months**

| | Follow-Up Month 1: Number and % of Patients Using Zantac Substitute | Follow-Up Month 2: Number and % of Patients Using Zantac Substitute | Follow-Up Month 3: Number and % of Patients Using Zantac Substitute |
|---|---|---|---|
| Treatment Group | 3<br>4% | 3<br>4% | 3<br>4% |
| Comparison Group | 2<br>2.0% | 2<br>2.0% | 1<br>1.0% |

The treatment group received the letter intervention in 9/94. The follow-up period for this group is 10/94-12/94. The comparison group's follow-up period is 1/95-3/95.

*Sources:* Maryland Medicaid, Mid-Atlantic DUR

The tables above present simple comparisons between treatment and comparison group Zantac users. The comparisons indicate that compared to the comparison group, the treatment group reduced their usage of Zantac in the three months after prescribers received intervention letters. These comparisons, however, do not account for the demographic differences between the treatment and comparison groups. Exhibit 6.4.6 below presents results from probit and OLS models that include demographic variables as covariates.

**Exhibit 6.4.6**

**HCFA DUR Demonstration Evaluation**
**Direct Effects of the Zantac Letter Intervention**

| | Dependent Variable | | |
|---|---|---|---|
| | Patient Refilled Zantac Prescription in Month 3 of Follow-Up Period | Patient Refilled Zantac Prescription in Month 3 of Follow-Up Period | Zantac Using Patients' Dosage in Month 3 of Follow-Up Period |
| | probit coefficient (se) | dF/dx evaluated at means of covariates | OLS regression coefficient (se) |
| white | -.084<br>(.217) | -.032<br>(.082) | -.066<br>(.088) |
| years of age | .004<br>(.007) | .001<br>(.003) | -.002<br>(.002) |
| female | .140<br>(.230) | .054<br>(.089) | .014<br>(.094) |
| baseline use of Zantac | .272***<br>(.094) | .104***<br>(.036) | -.057<br>(.041) |
| treatment group dummy variable | -.744***<br>(.244) | -.281***<br>(.089) | -.106<br>(.097) |
| constant | -.296<br>(.506) | ------ | 2.25**<br>(.210) |
| N | 177 | 177 | 104 |

The treatment group received the letter intervention in 9/94. The follow-up period for this group is 10/94-12/94. The comparison group's follow-up period is 1/95-3/95. Two patients with dosage levels above 19 pills per day were excluded from the analysis.

Marginal effects were calculated as follows: for continuous variables, dF/dx = f(xB) b(1) calculated at the means of X; for binary variables, the probabilities were calculated in a discrete fashion.

Baseline use of Zantac is measured by the number of Zantac prescriptions in the 3 months prior to the intervention (treatment group) or 3 months prior to being targeted for an intervention (comparison group).

The models displayed above confirm the comparison results presented earlier. As summarized in Exhibit 6.4.7 below, the probit models indicate that after controlling for race, sex, age and baseline Zantac use, treatment group members were significantly less likely than comparison group members to refill Zantac prescriptions during the third follow-up month. The magnitude of this statistically significant effect is meaningful from a policy perspective. Controlling for other factors, treatment group membership is associated with almost a .30 decrease in the probability of refilling a Zantac prescription in the third month of the follow-up period. The models also indicate that the intervention letters are not significantly associated with Zantac dosage decisions made during the third month of the follow-up period.

**Exhibit 6.4.7**

**HCFA DUR Demonstration Evaluation**
**Estimated Treatment Effects of RDUR Intervention**

| | |
|---|---|
| marginal effect of RDUR intervention on probability of refilling Zantac in third month of follow-up period | -.281***<br>(.089) |
| marginal effect of RDUR intervention on dosage level of Zantac users in third month of follow-up period | -.106<br>(.097) |

Estimated coefficients and standard errors are presented.
Marginal effects were calculated as follows: for continuous variables, dF/dx = f(xB) b(1) calculated at the means of X; for binary variables, the probabilities were calculated in a discrete fashion.

*Sources:* Maryland Medicaid, Mid-Atlantic DUR

### Results: Spillover Impact of the RDUR Letter Intervention

If the Zantac letter intervention was effective in making broad improvements in prescribing, one would expect that the impact of the letter intervention would "spill over" and affect the prescriptions of targeted prescibers' at-risk patients who were not named in the letters. The model displayed above was re-estimated using at-risk patients of targeted prescribers who were *not* named in the intervention letters as the treatment group. The comparison group was defined in the same way as before. Both the treatment and comparison groups were limited to patients who had been using Zantac for two consecutive months in the three months before the intervention (in the spill-over group) or before being targeted for an intervention (in the comparison group).

As summarized below in Exhibit 6.4.9, the models suggest that the effects of the Zantac letter intervention did indeed "spill-over" and affect at-risk patients whose physicians had received intervention letters but who were not specifically named in those letters. The estimated impact of the letter intervention on Zantac use in the spill-over population was almost as large as the estimated impact of the letter intervention on the patients named in the letters. Surprisingly, there also appears to be a spill-over effect on Zantac dosage, even though there was no evidence of a direct effect of the intervention on dosage. Spill-over patients had lower dosage levels of Zantac after the letter intervention compared to comparison group patients.

**Exhibit 6.4.8**

### HCFA DUR Demonstration Evaluation
### Spill-Over Effects of the Zantac Letter Intervention

| | Dependent Variable | | |
|---|---|---|---|
| | **Patient Refilled Zantac Prescription in Month 3 of Follow-Up Period** | **Patient Refilled Zantac Prescription in Month 3 of Follow-Up Period** | **Zantac Using Patients' Dosage in Month 3 of Follow-Up Period** |
| | **probit coefficient (se)** | **dF/dx evaluated at means of covariates** | **OLS coefficient (se)** |
| white | -.028<br>(.176) | -.011<br>(.068) | .012<br>(.080) |
| years of age | .005<br>(.005) | .002<br>(.002) | -.002<br>(.002) |
| female | .035<br>(.170) | .014<br>(.066) | .101<br>(.077) |
| baseline use of Zantac | .265*<br>(.149) | .103*<br>(.058) | -.120**<br>(.059) |
| spill-over group dummy variable | -.614***<br>(.193) | -.228***<br>(.066) | -.353***<br>(.079) |
| constant | .341<br>(.582) | ------ | 2.35***<br>(.246) |
| N | 282 | 282 | 169 |

Marginal effects were calculated as follows: for continuous variables, dF/dx = f(xB) b(1) calculated at the means of X; for binary variables, the probabilities were calculated in a discrete fashion.

Baseline use of Zantac is measured by the number of Zantac prescriptions in the 3 months prior to the intervention (treatment group) or 3 months prior to being targeted for an intervention (comparison group).

In dosage model, respondents who were not using Zantac and 2 respondents with dosage levels of more than 19 pills per day were excluded.

*Sources:* Maryland Medicaid claims , Mid-Atlantic DUR

**Exhibit 6.4.9**

**HCFA DUR Demonstration Evaluation**
**Estimated Spillover Effects of RDUR Intervention**

| | |
|---|---|
| marginal effect of RDUR intervention on probability of refilling Zantac in third month of follow-up period | -.281***<br>(.089) |
| marginal effect of RDUR intervention on dosage level of Zantac users in third month of follow-up period | -.106<br>(.097) |

Estimated coefficients and standard errors are presented.
Marginal effects were calculated as follows: for continuous variables, dF/dx = f(xB) b(1) calculated at the means of X; for binary variables, the probabilities were calculated in a discrete fashion.

*Sources:* Maryland Medicaid, Mid-Atlantic DUR

This evidence of spill-over effects, however, should be interpreted as suggestive evidence only. Because only limited information was available about the criteria that was used to target physicians for the Maryland letter intervention, it is not clear what factors led to spill-over patients not being named in the letters their physicians received. Clearly, without this information, one cannot be certain that the spill-over population is comparable to the comparison group population, who had been targeted for a Zantac letter intervention. The factors that led to spill-over patients not being named in the intervention letters may be related to later Zantac use and dosage. The results of this analysis, therefore, must be interpreted with caution.

## Limitations of the Analysis

The results of this analysis suggest that Maryland's H2 Blocker letter intervention was successful in changing prescriber behavior in the intended direction three months after the letters were sent. Providers who received letters were more likely to stop prescribing Zantac for their targeted patients compared to providers who did not receive letters. Although this result is consistent both with published studies and with the two other RDUR studies presented in this report, the analysis was limited by a lack of information about Maryland's RDUR program.

Very little data and information were available on the H2 Blocker letter intervention and on the Maryland Pharmacy Association's general RDUR procedures. In part because the RDUR program was discontinued in 1995, few knowledgeable people in Maryland were able or willing to discuss the program. This lack of information on the criteria used to select patients and physicians for the intervention limited the methods used to analyze spill-over effects. As mentioned earlier, one could not be certain that spill-over group patients were comparable to comparison group patients simply because it was not clear why spill-over patients had not been named in the intervention letters. For example, it is possible that the spill-over patients were not named in the intervention letters because their records suggested that they were likely to end Zantac use soon anyway. The comparison group, on the other hand, consisted of patients who were named in a later Zantac letter intervention. In this case, the spill-over effects results presented here might be spurious.

The spill-over analysis results, therefore, can only be interpreted as suggestive evidence that the impact of the RDUR letter intervention led to a broad change in prescribing behavior. Fortunately, the lack of information about the Maryland intervention was not as serious of a problem in the analysis of direct effects of the intervention because both the treatment and the comparison group had been targeted to be named in the intervention letters.

It also was not possible to separate the effects of the RDUR intervention from the effects of an academic detailing program that was being conducted simultaneously. Although it is not clear whether or not physicians who received intervention letters also received academic detailing, it is possible that at least a portion of the estimated letter intervention treatment effect actually should be attributed to academic detailing.

Even if one could be sure that the treatment and comparison physicians did not receive academic detailing, it is still not clear that the H2 Blocker letter intervention caused physicians to reassess their Zantac therapy decisions. Intervention letters were sent to targeted patients' physicians and pharmacists; pharmacists may have intervened instead of physicians. Furthermore, physicians in both the treatment and the comparison groups may have received many H2 Blocker RDUR letters over time for different patients. It is possible that part of the estimated treatment effect is actually a response to a series of letters instead a single letter intervention. It is also possible that the estimated treatment effect was diluted by contamination of the comparison group caused by previous letter interventions. Unfortunately, the available data does not allow an analysis of these topics.

Finally, this study focused on the effects of Maryland's H2 Blocker letter intervention three months after the letters were sent. It would be interesting to assess whether or not this effect lasted six months or even a year after the intervention. Although Medicaid claims for Zantac are available in these later months, the comparison group patients' prescribers received a H2 Blocker letter intervention during this time period. Because of the contamination of the comparison group, it is not reasonable to compare the treatment and comparison groups after the third follow-up month.

## Conclusions

The results of this analysis are consistent with published evaluations of DUR letter interventions targeted at H2 Blocker prescribing. The letter intervention, which was intended to reduce dosage and long-term use of H2 Blockers, is associated with a decrease in Zantac use during the follow-up period. The letters do not appear to have affected dosage decisions and use of Zantac substitutes. The results indicate that prescribers may have responded to the DUR intervention letter by stopping long-term Zantac therapy for their healthiest ulcer patients.

The analysis also uncovers suggestive evidence that the effects of the Zantac letter intervention extended to other at-risk patients who were not named in the letters. Spill-over patients were less likely than comparison group patients to be using Zantac during the follow-up period. Furthermore, spill-over patients were using lower dosage levels of Zantac relative to the comparison group. These results are not consistent with the findings of the two Iowa RDUR analyses presented in this report.

The Iowa studies suggested that the benefits of RDUR letters did *not* extend to at-risk patients who were not named in the letters.

As described earlier, it is possible that the limitations of the Maryland RDUR study design led to spurious evidence of spill-over effects. It is also possible, however, that the existence of spill-over effects depends on the nature of the RDUR intervention. The Maryland RDUR letter intervention was part of a broader educational effort that focused on reducing use of a widely used, expensive drug. The Iowa RDUR interventions, on the other hand, were intended to increase the use of drugs that had the potential to benefit patients but were not being used in many appropriate cases. The clinical context of these interventions may have affected their effectiveness, as well as the durability and scope of the letters' impact.

In light of the limitations discussed above, one cannot conclude with confidence that Maryland's RDUR letter intervention in September 1994 was successful in making lasting changes in prescribing. Instead, the results of this analysis offer evidence that either a single letter intervention, a series of letter interventions, or a combination of letter intervention and academic detailing is associated with a large decrease in Zantac prescribing in Maryland that potentially extended to patients who were not named in the letters. As state DUR boards continue to develop and evaluate their RDUR programs, this information is useful in understanding the benefits as well as the limitations of RDUR letter interventions.

# 7.0 Conclusions

## 7.1 Summary of Findings

Despite challenges faced in designing and implementing the demonstrations, both Project C.A.R.E. and the Iowa OPDUR Demonstration showed that pharmacies could be recruited into studies of prospective DUR with the knowledge that they would be randomized into treatment and control groups. In Washington, investigators gathered valuable information on the intensity and content of CS, showing that the most frequent interventions involve pharmacists counseling and educating patients, not interacting with prescribers or other pharmacists. Treatment pharmacies that received payment for CS also documented more CS than control pharmacies.

Iowa's demonstration generated useful information on the challenges of implementing and maintaining a well-validated OPDUR screening system. CS data collected in Iowa provided no direct evidence that the demonstration changed the behavior of pharmacists who received OPDUR messages, though it is clear that Iowa pharmacists had no particular incentive to keep timely or accurate CS, and there was some evidence that control group pharmacists may have been "overachieving" in documenting CS relative to treatment pharmacists. However the rate of claims reversals, which was higher for treatment than control pharmacies, provided Iowa investigators with some admittedly inconclusive evidence for an OPDUR effect on pharmacists' behavior.

This evaluation, however, found no evidence that prospective DUR, as implemented in the Project C.A.R.E. or the Iowa OPDUR Demonstration, had any measurable effects on the frequency of drug problems, the utilization and expenditures on prescription drugs and other medical services, or specific, well-defined clinical outcomes. In contrast, analyses of three retrospective drug use review interventions found statistically significant effects on prescription drug utilization in the hypothesized directions in each case.

## 7.2 Discussion

These findings do not prove that prospective drug use review is ineffective. Most pharmacies have, or will have, in-store computer based drug screening systems. With the spread of managed care and provider sponsored organizations, centralized, on-line screening will undoubtedly spread. Even if pharmacists weigh and evaluate messages they receive on potential drug problems against their own experience and information from other sources, a well-designed screening system is still valuable, as this evaluation found in surveys and focus groups with pharmacists. These demonstrations also showed that, with or without payment, pharmacists provide CS on a regular basis.

However, this evaluation did not find that the interventions in Iowa and Washington *added* to the underlying process of quality review that prescribers and pharmacists regularly use to screen drug prescribing in ways that would affect utilization, costs or outcomes. OPDUR may have provided information already available to pharmacists experienced with PC-based in-store screening systems. CS payment may not have been sufficient to support intensive interventions (pharmacists complained that the payment rate for "long" interventions, over six minutes, was inadequate). Project C.A.R.E.'s intervention was clearly not targeted toward detecting and correcting problems most likely to cause harm. Of course, there remains the question of whether or not this evaluation, given the relative infrequency of serious drug problems and given the number of pharmacies and patients in the two demonstrations (and, in some analyses, in the states added for the evaluation), could have found any effects. Had there been stronger evidence that the interventions affected pharmacists behavior as hypothesized, null findings in the utilization/cost and outcomes analyses would be cause for concern on statistical grounds. As it was, however, such evidence was lacking, a fact that is consistent with other findings in the study.

In contrast, retrospective drug use review seems relatively effective, but here too, caution would be advisable. This evaluation only reported on three interventions. The tests for effectiveness used in each of the three studies were necessarily based on quasi-experimental designs. Due to small sample sizes, no attempts were made to assess "downstream" effects of these interventions on other health services utilization and on clinical outcomes.

What do these findings mean for Medicaid, under mandate from OBRA 1990 to operate both prospective and retrospective drug use review programs? In one sense, the question of DUR policy has been rendered moot for a substantial proportion of Medicaid's client base, through the spread of managed care. State Medicaid directors can debate whether or not to require specific DUR models in contracts with managed care organizations, but the operational details will generally be left to the MCOs. Medicaid, however, remains the responsible fee-for-service payor for the most costly and vulnerable populations, the adult disabled and dually-eligible elderly. Consequently, policy-makers need to decide whether or not to fund further demonstrations and studies, to advocate changes in DUR requirements based on current knowledge, or to leave the present system in place.

The case for more research is supported by at least three arguments. First, the demonstrations evaluated here explored very new ground, and encountered problems in implementing and operating the interventions that were not fully resolved until the demonstrations had been under way for several months. On these grounds, perhaps, neither CS payment nor OPDUR was fairly tested. Second, based on feedback gained during the demonstration, designers of the interventions (especially Project C.A.R.E.) might have done things differently. For example, although the $4 payment for "brief" CS was viewed as fair by most pharmacists, the $6 payment for extended CS was not. Third, even on *a priori* grounds, the interventions tested may not have been the most effective models of prospective DUR. For example, it would have been instructive to asses the relative performance of Project C.A.R.E. and Iowa OPDUR against a third model that combined OPDUR capabilities with CS payment.

A less costly strategy might involve marginal policy changes, drawing on information learned in these studies about how pharmacists deliver CS and about effective methods of changing prescribing behavior. It was clear in both Iowa and Washington that, whatever incentives were introduced to encourage interventions to correct prescribing errors, pharmacists tended to spend a substantial amount of time counseling and educating patients. It was also clear in the three RDUR studies that prescribers responded to letters from state DUR boards targeted to specific drugs and patients. The literature also demonstrates the effectiveness, if not the cost-effectiveness, of more intensive interventions, like academic detailing. Since surveys and focus groups conducted in this evaluation depict an evolving but still distant relationship between pharmacists and prescribers, an effective use of Medicaid DUR funds might include both encouraging more RDUR and/or academic detailing interventions to address prescribing problems and a program for compensating pharmacists for counseling and educating patients.[1] This need not substitute for OPDUR, which can continue to support pharmacists' efforts to improve prescribing. It may, however, be a more productive use of public funds than generalized CS payment or attempts to standardize and validate OPDUR screening systems.

Finally, a decision to stay the present course is also reasonable. Nothing found in this evaluation strongly points toward DUR policy changes that will be cost effective for the Medicaid program. Increased understanding of the association of inappropriate drug utilization and adverse clinical outcomes would inform any improvements to DUR, but this requires basic research. More appropriately for state Medicaid, systems for DUR program performance monitoring could provide useful information to policy makers. Since OPDUR is (or will be) an operating program in most state Medicaid programs, it would be instructive to develop data capture systems that, at a minimum, monitor prescription drug utilization, in detail that matches the drug categories and screening criteria in operation. Also, standardized RDUR drug utilization measures that can be compared among interventions would provide state Boards and Medicaid with quantitative performance measures.

---

[1] One risk of this strategy, highlighted in focus groups conducted during this evaluation, is further exacerbation of prescriber/pharmacist relations. Several prescribers complained of pharmacists addressing prescribing problems first with the patient and then with the prescriber. Clear guidelines on reimbursable activities would be needed if such a targeted payment system were to be implemented.

# References

The AIRE Investigators. (1993). Effect of Ramipril on Mortality and Mobidity of Survivors of Acute Myocardial Infarction with Clinical Evidence of Heart Failure. *Lancet* 342, 821-348.

Alliance for Aging Research. (1998). When Medicine Hurts Instead of Helps: Preventing Medication problems in Older Persons. Washington DC.

American Medical Association, American Pharmaceutical Association, and Pharmaceutical Manufacturers Association (1991). *Principles of Drug Use Review (DUR).* Summer.

Andersen R, and Newman JF. (1973). Societal and Individual Determinants of Medical Care Utilization in the United States. *Milbank Memorial Fund Quarterly: Health & Society* 51(1), 95-124.

Applegate WB, et. al. (1983). Impact of a Cost-Containment Educational Program on Housestaff Ambulatory Clinic Charges. Medical Care 21(5), 486-496.

Armstrong E, and Denemark C. (1998). How Pharmacists Respond to On-Line, Real-Time Alerts. *J Am Pharm Assn* 38, 149-154.

Avron J, and Soumerai SB. (1983). Improving Drug-Therapy Decisions Through Educational Outreach. The New England Journal of Medicine 308(24), 1457-1463.

Avorn J, et. al. (1992) A Randomized Trial of a Program to Reduce the Use of Psychoactive Drugs in Nursing Homes. The New England Journal of Medicine 327(3), 168-173.

Baker DE. (1989). Misoprostol: New Drug Evaluation. *Practical Gastroenterology* 8, 8-14.

Barnes PJ, and Pedersen S. (1993). Efficacy and Safety of Inhaled Corticosteroids in Asthma. *American Review of Respiratory Disease* 148, S1-S26.

Bates DW, Spell N, and Cullen DJ. (1997) The Costs of Adverse Drug Events in Hospitalized Patients. *JAMA* 277, 307-11.

Baylis RD. (1994). Drug Utilization Review: A Description of Use for a Medicaid Population (Maryland) 1986-1994. *The Journal of Law, Medicine & Ethics* 22(3), 247-251.

Beasley R, et. al. (1991). Asthma Mortality and Inhaled Beta Agonist Therpay, *Aust NZ J Med* 21:753-763.

Beasley R, and Pearce NE. (1993). The Roles of Beta-Receptor Agonist Therapy in Asthma Mortality, Ann Arbor: CRC Press.

Berardi RR, and Dunn-Kucharski VA. (1993). Peptic Ulcer Disease: An Update. *American Pharmacy* NS3, 26-34.

Berardi RR, and Dunn-Kucharski VA. (1993). Peptic Ulcer Disease: An Update. *American Pharmacy* NS33, 26-34.

Bootman JL, Harrison DL, and Cox E. (1997). The Health Care Cost of Drug-Related Morbidity and Mortality in Nursing Facilities. *Archives of Internal Medicine* 157, 2089-2096.

Borghi C, and Ambrosioni E. (1996). A Risk-Benefit Assessment of ACE Inhibitor Therapy Post-Myocardial Infarction. *Drug Safety* 14, 277-287.

Brasfield, K.H. (1991). Practical Psychopharmacologic Considerations in Depression. *Nursing Clinics of North America* 26, 651-63.

Braun S, Boyko V, Behar S, Reicher-Reiss H, Shotan A, Schlesinger Z, Rosenfeld T, Palant A, Friedenshon A, Laniado S and Goldbourt U. (1996). Calcium Antagonists and Mortality in Patients with Coronary Artery Disease: A Chohort Study of 11,575 Patients. *American College Cardiology* 28, 7-11.

Britton ML, and Lurvey PL. (1991). Impact of Medication Profile Review on Prescribing in a General Medicine Clinic. American Journal of Hospital Pharmacy 48, 265-270.

Brouwer RM, Follath F, and Buhler FR. (1985). Cardiovascular Adversity of the Calcium Antagonist Beta-Blocker Combination: Implications for Antihypertensive Therapy. *Journal of Cardiovascular Pharmacology* 7:S38-S44.

Brown DJ, Helling DK, and Jones ME. (1979). Evaluation of Clinical Pharmacist Consultations in a Family Practice Office. American Journal of Hospital Pharmacy 36, 912-915.

Brufsky JW, et. al. (1998). Shifting Physician Prescribing to a Preferred Histamine-2-Receptor Antagonist. *Medical Care* 36(3), 321-331.

Brymer, C., Winograd, C.H. (1992). Fluoxetine in Elderly Patients: is there cause for Concern? *J Amer Geriatric Soc* 40, 902-905.

Buring JE, Glynn RJ, and Hennekens CH. (1995),. Calcium Channel Blockers and Myocardial Infarction: A Hypothesis Formulated but Not Yet Tested. *JAMA* 274, 654-655.

Burkhardt, D., Raeder, E., Muller, V., Imhof, P., and Newbauer, H. (1978). Cardiovascular Effects of Tricyclic and Tetracyclic Antidepressants. *JAMA* 239, 213-216.

Byrns PJ. (1993). Drug utilization review criteria. Paper presented at the Workshop on Innovation in Medicaid Drug Utilization Review (DUR). Chantilly, VA.

Campbell AJ, Borrie MJ, and Spears GF. (1989). Risk Factors for Falls in a Community-based Prospective Study of People 70 Years and Older. *Journal of Gerontology* 44:M112-M117.

Carson JL, et. al. (1987). The Relative Gastrointestinal Toxicity of the Nonsteroidal Anti-inflammatory Drugs. *Archives of Internal Medicine* 147(6), 1054-9.

Carter BL , et. al. (1984). Evaluation of Family Physician Prescribing: Influence of the Clinical Pharmacist. Drug Intelligence and Clinical Pharmacy 18, 817-821.

Cehli R, ed. (1989). Treatment and Prevention of NSAID-Induced Gastropathy. Royal Society of Medicine Services International Congress and Symposium Series 147.

Center for Disease Control, Vital and Health Statistics. (1997). Series 10: Data from the National Health Survey 194, 1-89.

Center on Drugs and Public Policy. (1996). State of Maryland Drug Use Review Board Screening Criteria 1995-1996 Baltimore: University of Maryland at Baltimore School of Pharmacy.

Chrischilles EA, Helling DK, and Aschoff CR. (1989). Effect of Clinical Pharmacy Services on the Quality of Family Practice Physician Prescribing and Medication Costs. DICP 23, 333-341.

Chrischilles EA, et. al. (1997). The Iowa Medicaid OPDUR Demonstration Project Annual Report. Des Moines: Iowa Department of Human Services, State of Iowa; and Iowa City: Department of preventive Medicine and Environmental Health, College of Medicine, The University of Iowa under HCFA Cooperative Agreement No. 11-C-90232/7.

Christensen D, Fasset W, and Andrews A. (1994). *American Pharmacy.*

Chutka, D.S. (1990) Cardiovascular Effects of the Antidepressants: Recognition and Control. *Geriatrics* 45, 55-63.

Clauzel AM. (1994). Inhaled β-2 Agonists and Steroids, *Clin Rev in Allergy* 12, 43-64.

Coccaro, E.F., and Siever, L.J. Second Generation Antidepressants: A Comparative Review. *J Clin Pharmocol* 25, 241-260.

Collins TM, et. al. (1997). A Controlled Letter Intervention to Change Prescribing Behavior: Results of a Dual-Targeted Approach. *Health Services Research* 32(4), 471-489.

Committee on Therapeutic Education of the Medical and Surgical Faculty of Maryland and the University of Maryland School of Pharmacy. (1993). *Equal Time* 2(2).

Conti CR. (1996). Re-examining the Clinical Safety and Roles of Calcium Antagonists in Cardiovascular Medicine. *American Journal of Cardiology* 78, 13-18.

Cookson, J. (1993). Side Effects of Antidepressants. *Br J Psychiatr* 20, 20-24.

Cumming RG. (1998). Epidemiology of Medication-related Falls and Fractures in the Elderly. *Drugs Aging* 12, 43-53.

Cumming RG, and Klineberg RJ (1994). Fall Frequency and Characteristics and the Risk of Hip Fractures. *Journal of the American Geriatrics Society* 42, 774-778.

Cunningham, L.A. (1994). Depression in the Medically Ill: Choosing an Antidepresant. *J Clin Psychiatry* 55(9)(Suppl A), 90-97.

The Danish Study Group on Verapamil in Acute Myocardial Infarction. (1986). The Danish Studies on Acute Myocardial Infarction. *Br Journal of Clinical Pharmocology* 21, 197S-204S.

Davis DA, et. al. (1995). Changing Physician Performance: A Systematic Review of the Effect of Continuing Medical Education Strategies. JAMA 274(9), 700-705.

Davis DA, et. al. (1992). Evidence for the Effectiveness of CME: A Review of 50 Randomized Controlled Trials. JAMA 268(9), 1111-1117.

Department of Health and Human Services (1991). International Classification of Diseases, 9th Revision, Publication No. (PHS) 91-1260.

Department of Health and Mental Hygiene. (1994) State of Maryland Drug Use Review Board: Annual Report, Federal Fiscal year 1993. MD: Department of Health and Mental Hygiene.

Department of Health and Mental Hygiene. (1998). State of Maryland Drug Use Review Board: Annual Report, Federal Fiscal year 1997. MD: Department of Health and Mental Hygiene.

Diwan V, et. al. (1995). Effects of "Group Detailing" on the Prescribing of Lipid-Lowering Drugs: A Randomized Controlled Trial in Swedish Primary Care. Journal of Clinical Epidemiology 48(5), 705-711.

Dobie III RL, and Rascati KL. (1994). Documenting the Value of Pharmacist Interventions. *American Pharmacy* NS34(5), 50-54.

Edelson JT, Tosteson AN, and Sax P. (1990). Cost-effectiveness of Misoprostol for Prophylaxis Against Nonsteroidal Anti-inflammatory Drug-Induced Gastrointestinal Tract Bleeding. *JAMA* 264, 41-47.

Edwards, J.G. (1992). Selective Serotonin Reuptake Inhibitors. *Br Med J* 304, 1644-6.

Eisenberg JM. (1985). Physician Utilization: The State of Research About Physicians' Prescribing Patterns. Medical Care 23(5), 461-483.

Eisenberg JM, and Williams SV. (1981). Cost Containment and Changing Physicians' Practice Behavior. JAMA 246(19), 2195-2201.

Epstein, M. (1995). Calcium Antagoinists: Should Continue to Be Used for First-Line Treatment of Hypertension? *Arch Intern Med* 155, 2150-2156.

Evans JH, Hwang Y, and Nagarajan N. (1995). Physicians' Response to Length-of-Stay Profiling. Medical Care 33(11), 1106-1119.

Fagan TC. (1995). Calcium Antagonist and Mortality: Another Case of the need for Clinical Judgment. *Archives of Internal Medicine* 155, 2145.

Falk RH, and Leavitt JI. (1991). Digoxin for Atrial Fibrillation: A Drug Whose Time Has Gone. *Annals of Internal Medicine* 114, 573-575.

Feighner, J.P. (1994). The Role of Venlafaxine in Rational Antidepressant Therapy. *J Clin Psychiatry* 55(9)(Suppl), 62-70.

Finkel, S.I. Efficacy and Tolerabiltiy of Antidepressant Therapy in the Old-Old. *J Clin Psychiatry* 57 (Suppl 5), 23-28.

Forstrom MJ, et. al. (1990) Title. DICP 24, 304-309.

Frazier LM, et. al. (1991). Can Physician Education Lower the Cost of Prescription Drugs? Annals of Internal Medicine 115,116-121.

Freis ED, and Papademetrious V. (1996). Current Drug Treatment and Treatment Patterns with Antihypertensive Drugs. *Drugs* 52, 1-16.

Fries JF, et. al. (1989). Management of NSAID-Induced Gastric Mucosal Injury. *Postgraduate Perspectives* 1,1-20.

Fries JF, et. al. (1991). Nonsteroidal Anti-Inflammatory Drug-Associated Gastropathy: Incidence and Risk Factor Models [see comments]. Department of Medicine, Stanford University School of Medicine, California. *American Journal of Medicine* 91(3), 213-22 (Comment in: *Am J Med* (1991) 91(3), 209-12).

Furberg CD, Psaty BM, and Meyer JV. (1995). Nifedipine: Dose-Related Increase in Mortality in Patients with Coronary Heart Disease. *Circulation* 92, 1326-1331.

Gabriel SE, Jaakkimainen RL, and Bombardier C. (1993). The Cost-effectiveness of Misoprostol for Nonsteroidal Anti-inflammatory Drug-Associated Adverse Gastrointestinal Events. *Arthritis and Rheumatism* 36, 447-59.

Giulekas D, et. al. Influence of Pindolol on Asthmatics and Effect of Bronchodilators. (1986). *Respir* 50, 158-166.

Glasser SP, et. al. (1991). Angina to Placebo Periods in Drug Trials. *JAMA* 265, 1550-1554.

Glassman, A.H. (1984). Cardiovascular Effects of Tricyclic Antidepressants. *Ann Rev Med.* 35, 503-11.

Glassman AH, Bigger JT Jr., Giardina EV, Kantor SJ, Perel JM, Davies M. (1979). Clinical Characteristics of Imipramine-Induced Orthostatic Hypotension. *Lancet* 1, 468-72.

Glassman AH, Johnson LL, Giardina EGV, Walsh BT, Toose SP, Cooper TB and Bigger JT. (1983). The Use of Imipramine in Depressed Patients with Congestive Heart Failure. *JAMA* 250, 1997.

Glassman, A.H., Preud'homme, X.A. (1993). Review of the Cardiovascular Effects of Heterocyclic Antidepressants. *J Clin Psychiatry* 54 (Suppl), 16-22.

Glassman, A.H., Roose, S.P. (1994). Risks of Antidepressants in the Elderly: Tricyclic Antidepressants and Arrhythmia - Revising Risks. *Gerontol* 40 (Suppl 1), 15-20.

Glassman, A.H., Roose, S.P., Bigger, J.T. Jr. (1993). The Safety of Tricyclic Antidepressants in Cardiac Patients. Risk-benefit Reconsidered. *JAMA* 1993. 269(20), 2673-5.

Goldman LS, Alexander RC, Luchins DJ. (1986). Monoamine Oxidase Inhibitors and Tricyclic Antidepressants: Comparison of their Cardiovascular Effects. *J Clin Psychiatry.* 47, 225-229.

Goldstein RE, Boccuzzi SJ, Cruess D, Nattel S. (1991). Diltiazem Increases Late-Onset Congestive Heart Failure in Postinfarction Patients with Early Reduction in Ejection Fraction. *Circulation* 83, 52-60.

Gomes JA, et. al. (1993). Double-blind Comparison of Efficacy and Gastroduodenal Safety of Diclofenac/Misoprostol, Piroxicam and Naproxen in the Treatment of Osteoarthritis. *Annals of Rheumatic Diseases* 52, 881-5.

Gottlieb GL. (1990). Sleep Disorders and their Management: Special Considerations in the Elderly. American Journal of Medicine 88, 29S-33S.

Grad RM. (1995). Benzodiazepines for Insomnia in Community-Dwelling Elderly: A Review of Benefit and Risk. *Journal of Family Practice* 41,473-481.

Greco PJ, and Eisenberg JM. (1993). Changing Physicians' Practices. The New England Journal of Medicine 329(17),1271-1274.

Griffin MR, et. al. (1991). Nonsteroidal Anti-inflammatory Drug Use and Increased Risk for Peptic Ulcer disease in Elderly Persons. Vanderbilt University School of Medicine, Nashville, Tennessee. *Annals of Internal Medicine* 114(4), 257-63.

Griffin MR, Ray WA, and Schaffner W. (1988). Nonsteroidal Anti-inflammatory Drug Use and Death from Peptic Ulcer in Elderly Persons. Vanderbilt University School of Medicine, Nashville, Tennessee. *Annals of Internal Medicine* 109(5), 359-63.

Guo JJ, et. al. (1995). Retrospective Drug Utilization Review and the Behavior of Medicaid Prescribers: An Empirical Marginal Analysis. *Clinical Therapeutics* 17,1174-1187.

Gurwitz JH, Soumerai SB, and Avorn J. (1990). Improving Medication Prescribing and Utilization in the Nursing Home. Journal of the American Geriatric Society 38(5),542-551.

Gutierrez G, et. al. (1994). Changing Physician Prescribing Patterns: Evaluation of an Educational Strategy for Acute Diarrhea in Mexico City. Medical Care 32(5), 436-446.

Gabriel SE, Jaakkimainen RL, and Bombardier C. (1993). The Cost-effectiveness of Misoprostol for Nonsteroidal Anti-inflammatory Drug-Associated Adverse Gastrointestinal Events. *Arthritis and Rheumatism* 36, 447-59.

Giulekas D, et. al. Influence of Pindolol on Asthmatics and Effect of Bronchodilators. (1986). *Respir* 50, 158-166.

Halper JP, Mann JJ. (1988). Cardiovascular Effects of Anti depressant Medication. *Br J Psychiatry* 153(Suppl 3), 87-98.

Hamann SR, Kaltenborn KE, McAllister RG. (1987). Hemodynamic and Pharmacokinetic Aspects of the Interactions Between Verapamil and Pindolol. *Pharmacology* 36, 326-336.

Hamann SR, Kaltenborn KE, Vore M, Tan TG, McAllister RG. (1985). Cardiovascular and Pharmacokinetic Consequences of Combined Administration of Verapamil and Proprandolol in Dogs. *American Journal of Cardiology* 56, 147-156.

Hansson L, Zanchetti A, HOT Study Group. (1994). The Hypertension Optimal Treatment (HOT) Study: Patient Characteristics: Randomization, Risk Profiles, and Early Blood Pressure Results. *Blood Pressure* 3, 322-327.

Hawkey CJ, et. al. (1990). Effect of Reactive Pharmacy Intervention on Quality of Hospital Prescribing. British Medical Journal 30, 986-990.

Hawkey CJ. (1993). Healing and Prevention of NSAID-Induced Peptic Ulcers. *Scandinavian Journal of Gastroenterol*ogy 29 (suppl 201,42-4.

Hayes, JR, Born GP, Rosenbaum AH. (1977). Incidence of Orthostatic Hypotension in Patients with Primary Affective Disorders Treated with Tricyclic Antidepressants. *Mayo Clin Proc* 52, 509-12.

Hepler CD, and Strand L. (1990). Opportunities and Responsibilities in Pharmaceutical Care. *American Journal of Hospital Pharmacy* 47, 533-543.

Herfindal ET, Bernstein LR, and Kishi DT. (1985). Impact of Clinical Pharmacy Services on Prescribing in a Cardiothoracic/Vascular Surgical Unit. Drug Intelligence and Clinical Pharmacy 19, 440-444.

Herings RM, et. al. (1995). Benzodiazepines and the Rrisk of Falling Leading to Femur Fractures: Dosage more Important than Elimination Half-life. *Archives of Internal Medicine* 155, 1801-1807.

Hillman AL, Pauly MV, and Kerstein JJ. (1989). How Do Financial Incentives Affect Physicians' Clinical Decisions and the Financial Performance of Health Maintenance Organizations? The New England Journal of Medicine 321(2), 86-92.

Hillman AL, and Bloom BS. (1989). Economic Effects of Prophylactic Use of Misoprostol to Prevent Gastric Ulcer in Patients Taking Nonsteroidal Anti-inflammatory Drugs. *Archives of Internal Medicine* 149, 2061-65.

Hofer, CA. (1993). Verapamil Intoxification: A Literature Review of Overdoses and Discussion of Therapeutic Options. *American Journal of Medicine* 95, 431-438.

Horton R. (1995). Spinning the Risks and Benefits of Clacium Antagonists. *Lancet* 346, 586-587.

Howes LG. (1995). Which Drugs Affect Potassium? *Drug Safety* 12, 240-244.

Hueston WJ, et. al. (1997). Evaluation and Treatment of Respiratory Infections: Does Managed Care Make a Difference? The Journal of Family Practice 44(6): 572-577.

Huszonek JJ, Dewan MJ, Koss M, Hardoby WJ, and Ispahani A. (1993). Antidepressant Side Effects and Physician Prescribing Patterns. *Ann Clin Psychiatr* 5, 7-11.

ISIS-4 Collaborative Group. (1995). ISIS-4: A Randomised Factiorial Trial Assessing Early Oral Captopril, Oral Mononitrate, and Intravenous Magnesium Suplhate in 58050 Patients with Suspected Acute Myocardial. *Lancet* 345, 669-85.

Jackson WK, Roose SP, Glassman AH. (1987). Cardiovascular Toxicity and Tricyclic Antidepressants. *Biommed Pharmacother* 41, 377-382.

Jameson J, VanNoord G, and Vanderwoud K. (1995). The Impact of a Pharmacotherapy Consultation on the Cost and Outcome of Medical Therapy. *The Journal of Family Practice* 41(5), 469-472.

Johnson, JA. and Bootman, JL. (1995). Drug-Related Morbidity and Mortality: A Cost-of-Illness Model. Archives of Internal Medicine 155, 1949-56.

Johnson JA, and Bootman JL.(1995). Drug-Related Morbidity and Mortality: A Cost of Illness Model. *Archives of Internal Medicine* 155L19, 49-56.

Johnston DL, Lesoway R, Humen DP, et al. (1985). Clinical and Hemodynamic Evaluation of Propranolol in Bombination with Verapamil, Nifedipine and Diltiazem in Extertional Angina Pectoris: A Placebo-Controlled, Double-Blind, Randomized Crossover Study. *American Journal of Cardiology* 55, 680-687.

Kloner RA. (1995). Nifedipine in Ischemic Heart Disease. *Circulation* 92, 1074-1078.

Knapp D, and Erwin G. (1992). Screening Criteria for Outpatient Drug Use Review. Final Report to HCFA. Baltimore: University of Maryland School of Pharmacy.

Knapp DA, and Colleagues at the Center on Drugs and Public Policy. (1989). Recommendations for Developing and Implementing an Outpatient Drug Use Review Program Under the Medicare Catastrophic Coverage Act. Special report to the Health Care Financing Administration. MD: University of Maryland.

Knapp DA, and Colleagues at the Center on Drugs and Public Policy. (1989). Drug Utilization Review for the Ambulatory Elderly: A Review of the Research Literature. Special report to the Health Care Financing Administration. MD: University of Maryland.

Koher L, Torp-Pedersen C, Carlson JE, Trandolapril Cardiac Evaluation (TRACE) Study Group, et al. (1995). A Clinical Trial of the Angiotensin-Covering Enzyme Inhibitor Trandolapril in Patients with Left Ventricular Dysfunction after Myocardial Infarction. *New England Journal of Medicine* 333, 1670-6.

Kojima S, et. al. (1995). Comparison of First-Line Antihypertensive Drugs by a Randomized Cross-over Method--A Preliminary Report. *Hypertens Rev* 18, 303-305.

Kopera H. (1978). Anticholinergic and Blood Pressure Effects of Mianserin, Amitriptline and Pacebo. *Br J Clin Pharmacol* 5(Suppl 1), 29s-34s.

Keys PW, et. al. (1995). Computer Guided Academic Detailing as Part of a Drug Benefit Program. American Journal of Health Systems and Pharmacy 52, 2199-2204.

Kralewski JE. (1990). Drug Utilization Review in the Private Sector. Final Report of a study funded by the Health Care Financing Administration. Cooperative Agreement No. 99-C-99169/5-01 T2.

Krebs R .(1983). Adverse Reactions with Calcium Antagonists. *Hypertension* 5(Part2):II125-129.

Laird, L.K., Lydiard, R.B., Morton, W.A., Steele, T.E., Kellner, C., Thompson, N.M., and Ballenger, J.C. (1993). Cardiovascular Effects of Imipramine, Fluvoxamine, and Placebo in Depressed Outpatients. *J Clin Psychiatr* 54, 224-228.

Lang DM, Sherman MS, and Polansky M. (1997). Guidelines and Realities of Asthma Management. *Archives of Internal Medicine*.

Lazarou J, Pomeranz BH, and Corey, PN. (1998). Incidence of Adverse Drug Reactions in Hospitalized Patients: A Meta-analysis of Prospective Studies. *JAMA* 279,1200-05.

Lenfent C. (1995). The Calcium Channel Blocker Scare: Lessons for the Future. *Circulation* 92, 1068-1073.

Leonard, B.E. (1993). The Comparative Pharmacology of New Antidepressants. *J Clin Psychiatry* 54(8 Suppl), 3-15.

Lipton HL and Bird JA. (1991). Drug Utilization Review: State of the Art From an Academic Perspective. *Clinical Pharmacology and Therapeutics* 50, 616-619.

Lin EH, et. al. (1997) Achieving Guidelines for the Treatment of Depression in Primary Care: Is Physician Education Enough? Medical Care 35(8), 831-842.

Lipton HL, and Bird JA. (1993). Drug Utilization Review in Ambulatory Settings: State of the Science and Directions for Outcomes Research. Medical Care 31(12), 1069-1082.

Lobeck F, Traxler WT, and Bobinet DD. (1989). The Cost Effectiveness of a Clinical Pharmacy Service in a Outpatient Mental Health Clinic. Hospital and Community Psychiatry 40(6), 643-645.

Lord SR. (1990). Falls in the Elderly: Admissions, Bed Use, Outcome and rojections. *Med J Aust* 153(2), 117-118.

Lundborg CS, Hensjo LO, and Gustafsson L. (1997). "Academic Drug-Detailing": From Project to Practice in a Swedish Urban Area. European Journal of Clinical Pharmacology 52, 167-172.

Materson BJ, and Preston RA. (1994). Angiotensin-Converting Enzyme Inhibitors in Hypertension; A Dozen Years of Experience. *Archives of Internal Medicine* 154, 513-523.

McFadden ER, and Gilbert IA. (1992). Asthma. *New England Journal of Medicine* 327(27) 1928-1937.

McMahon GF. (1990). Calcium Channel Blockers. In Management of Essential Hypertension. Mt. Kisco: Futura Publishing.

Messerli FH. (1995). Case-Control Study, Meta-Analysis, and Bouillabaisse: Putting the Calcium Scare into Context. *Annals of Internal Medicine* 123, 888-889.

Miller MJ,. and Ortmeier BG. (1995). Factors Influencing the Delivery of Pharmacy Services. *American Pharmacy* NS35(1):39-45.

Monane M, and Avorn J.(1996). Medications and Falls: Causation, Correlation, and Prevention. *Clinical Geriatric Medicine* 12, 847-58.

Montgomery, S.A., (1993). Symposium Chariperson. Venlafaxine: A New dimension in Antidepressant Pharmacotherapy. *J Clin Psychiatry*. 54(3), 119-126.

Morrow RW, Gooding AD, and Clark C. (1995). Improving Physicians' Preventive Health Care Behavior Through Peer Review and Financial Incentives. Archives of Family Medicine 4, 165-169.

Mullen, PB. (1990). Prescription Drugs Illinois: Publications International, Ltd.

Muller OF, Goodman N, Bellet S. (1961). The Hypotensive Effect of Imipramine Hydroloride in Patients with Cardiovascular Disease. *Clin Pharmoacol Ther* 2, 300-7.

Neergaard L. (1996). Most Hypertension Drugs Deemed Safe by Scientific Panel. New York: Associated Press.

Nelson, J.C. (1993). Combined Treatment Stategies in Psychiatry. *J Clin Psychiatry* 54, 42-49.

The NORDIL Group. (1993). The Nordic Diltiazem Study (NORDIL): A Prospective Intervention Trial of Calcium Antagonist Therapy in Hypertension. 2, 314-321.

O'Loughlin JL, et. al. (1993). Incidence of, and Risk Factors for, Falls and Injurious Falls Among the Community-Dwelling Elderly. *American Journal of Epidemiology* 137, 342-354.

Opie LH. (1988). Calcium Channel Antagonists. Part IV: Side Effects and Contraindications Drug Interactions and Combinations. *Cardiovascular Drugs and Therapy* 2, 177-189.

Opie LH, and Messerli FH. (1995). Nifedipine and Mortality: Grave Effects in the Dossier. *Circulation* 92. 1068-1073.

Packer M, Meller J, Medina N, et al. (1982). Hemodynamic Consequences of Combined Beta-Adrenergic and Slow Calcium Channel Blockade in Man. *Circulation* 65, 660-668.

Pahor M, Guralnik JM, Corti M-C, Foley DJ, Carbonin P, Havlik RJ. (1995). Long-term Survival and Use of Antihypertensive Medications in Older Persons. *Journal of American Geriatric Society* 43, 1191-1197.

Patakas D, et. al. (1983). Beta-blockers in Bronchial Asthma: Effect of Propranolol and Pindolol on Large and Small Airways. *Thorax* 38, 108-112.

Pearigen PD, and Benowitz NL. (1991). Poisoning due to Calcium Antagonists: Experience with Verapamil, Diltiazem and Nifedipine. *Drug Safety* 6, 408-430.

Pedersen KE. (1985). Digoxin Interactions: The Influence of Quindine and Verapamil on the Pharmacokinetics and Receptor Binding of Digitalis Glycosides. *Acta Med Scand* (Suppl 697), 12-40.

Pepine CJ. (1996(. Ongoing Clinical Trials of Angiotensin-Converting Enzyme Inhibitors for Treatment of Coronary Artery Disease in Patients With Preserved Left Ventricular Function. *American College of Cardiology* 27, 1048-1052.

Peterson CD, and Goldberg DE. (1989). Pharmacy-Coordinated Process for Evaluating Physician Drug Prescribing. American Journal of Hospital Pharmacy 46, 1787-1791.

Peterson GM, and Sugden JE. (1995). Educational Program to Improve the Dosage Prescribing of Allopurinol. The Medical Journal of Australia 162, 74-77.

Pfeffer MA, Braunwald E, Moye LA, et al. (1992). Effect of Captoril on Mortality and Morbidity in Patients with Left Ventricular Dysfunction after Myocardial Infarction. *New England Journal of Medicine* 327, 669-77.

Peterson GM, et. al. (1996). Evaluation of Physician Intervention Letters. *Medical Care* 34, 760-766.

Pinter JK, Wiley MN, and Pennypacker L. (1994). Prevention of NSAID-Induced Gastropathy in the Elderly. *The Consultant Pharmacist* 9, 568-579.

Preskorn, S.H. (1994). Antidepressant Drug Selection: Criteria and Options. *J Clin Psychiatry* 55, 9(Suppl A), 6-22.

Preskorn, S.H. (1993). Pharmacokinetics of Antidepressants: Why and How They are Relevant to Treatment. *J Clinical Psychiatry* 54, 14-34.

Preskorn, S.H. (1996). Reducing the Risk of Drug-Drug Interactions: A Goal of Rational Drug Development. *J Clinical Psychiatry* 57(Suppl 1), 3-6.

Preskorn, S.H., Irwin, H.A. (1982). Toxicity of Tricyclic Antidepressants - Kinetics, Mechanisms, Interventions: A Review. *J Clinical Psychiatry* 43, 151-156.

Prichard J. (1986). Beta-blockers and the Patient with Bronchial Disease, *Irsih Medical Journal* 79(11), 303-304.

Psaty BM, Heckbert SR, Koepsell TD, Siscovick DS, Raghunathan TE, Weiss NS, Rosendaal FR, Lemaitre RN, Smith NL, Wahl PW, Wagner EH, and Furberg, CD. (1995). The Risk of Myocardial Infarction Associated with Antihypertensive Drug Therapies. *JAMA* 274, 620-625.

Rakel, RE. (1996). Ed. Conn's Current Therapy. W.B. Saunders Company.

Ramoska EA, Spiller HA, Winter M and Borys D. (1993). A One Year Evaluation of Calcium Channel Blocker Overdose: Toxicity and Treatment. *Ann Emer Med* 22, 196-200.

Rascati KL, Okano GJ, and Burch C. (1996). Evaluation of Physician Intervention Letters. *Medical Care* 34, 760-766.

Ray W.A. and Griffin M.R. (1991). Cyclic Antidepressants and the Risk of Hip Fracture. *Arch Intern Med* 151, 754-756.

Ray WA, Griffin MR, Schaffner W, Baugh D.K., and Melton III LJ. (1987). Psychotropic Drug Use and the Risk of Hip Fracture. *New England Medical Journal* 316, 363-9.

Ray WA, et. al. (1985). Improving Antibiotic Prescribing in Outpatient Practice. Medical Care 23(11), 1307-1313.

Ray WA, Schaffner W, and Federspiel CF. (1985). Persistence of Improvement in Antibiotic Prescribing in Office Practice. JAMA 253(12), 1774-1776.

Ray WA, et. al. (1986). Reducing Long-Term Diazepam Prescribing in Office Practice: A Controlled Trial of Educational Visits. JAMA 256, 2536-2539.

Ray WA, et. al. (1987). Reducing Antipsychotic Drug Prescribing for Nursing Home Patients: A Controlled Trial of the Effect of an Educational Visit. American Journal of Public Health 77, 1448-1450.

Ray WA, Griffin MR,and Downey W. (1989). Benzodiazepines of Long and Short Elimination Half-life and the Risk of Hip fracture. *JAMA* 262, 3303-3307.

Ray WA, Griffin MR, and Malcolm E. (1991). Cyclic antidepressants and the risk of hip fracture. *Archives of Internal Medicine* 15,754-756.

Ray WA, et. al. (1987). Psychotropic Drug Use and the Risk of Hip Fracture. *New England Journal of Medicine* 316, 363-369.

Ray WA, et. al. (1990). Incidence of Hip Fracture in Saskatchewan, Canada, 1976-1985. *American Journal of Epidemiology* 131,502-509.

Rengo F, Carbonin P, Pahor M, De Caprio L, Bernanbei R, Ferrara N, Carosella L, Acanfora D, Parlati S, Vitale D, and the CRIS Investigators. (1996). A Controlled Trial of verapamil in Patients After Acute Mycocardial Infartion: Results of the Calcium Antagonist Reinfarction Study (CRIS). *American Journal of Cardiology* 77, 366-369.

Renhnquist N, Hjendahl P, Billings E, Bjorkardur I, Eriksson SV, Forslund L, Held C, Nasman P, Wallen NH. (1996). Effects of Metoprolol verses Verapamil in Patients with Stable Angina Pectoris: The Angina Prognosis Study in Stockholm (APSIS). *Eur Heart Journal* 17, 76-78.

Rogers B. (1993). Regression Standard Errors in Clustered Samples. In: *STATA Technical Bulletin Reprints*, Vol. 3 (1994). College Station, TX: STATA Press.

Roose S.P., Dalack GW. (1992). Treating the Depressed Patient with Cardiovascular Problems. *J Clincial Psychiatry* 53(9 Suppl), 25-31.

Roose, S.P., Glassman, A.H. (1994). Antidepressant Choice in the Patient with Cardiac Disease: Lessons for the Cardiac Arrhythmia Suppression Trial (CAST). *J Clinical Journal* 55(Suppl A), 83-89.

Roose, S.P., Glassman, A.H., Giardina, E.G.V., Walsh, B.T., Woodring, S., and Bigger, J.T. (1987). Tricyclic Antidepressants in Depressed Patients with Cardiac Conduction Disease. *Arch Gen Psychiatry* 44, 273-275.

Roose, S.P., Glassman, A.H., Siris S.G., Walsh B.T., Bruno, R.L., Wright L.B. (1981). Comparison of Imipramine - and Nortriptyline-induced Orthostatic Hyptension: A Meaningful Difference. *J Clinical Psychopharmacol.* 1, 316-19.

Ruderdorfer, M.V., Potter, W.Z. (1989). Antidepressants:A Comparative Review of the Clincial Pharmacology and Therapeutic Use of the "Newer" Verses the "Older" Drugs. *Drugs* 37, 713-738.

Rupp, MT. (1991). *American Pharmacy* N332-360.

Rupp MT. (1992). Value of Community Pharmacists' Interventions to Correct Prescribing Errors. *The Annals of Pharmacotherapy* 26,1580-84.

Rupp MT. (1988). Evaluation of Prescribing Errors and Pharmacist Interventions in Community Practice: An Estimate of 'Value Added'. *American Pharmacy* NS28(12), 766-770.

Rupp MT, et. al. (1988). Documenting Prescribing Errors and Pharmacist Interventions in Community Practice. *American Pharmacy* NS28(9), 574-580.

Rupp MT, DeYoung M, and Schondelmeyer SW. (1992). Prescribing Problems and Pharmacist Interventions in Community Practice. *Medical Care* 30(10), 926-940.

Russell RP. (1988). Side Effects of Calcium Channel Blockers. *Hypertension* 11, 1142-1144.

Ruzicka M, and Leenen FH. (1995). Relevance of Internittent Increases in Sympathetic Activity for Adverse Outocme on Short-Acting Calcium Antagonists. In Hypertension: Pathology, Diagnosis, and Management. Edition II. (Eds.) Laragh JH, and Brenner BM. New York: Raven Press.

Sattin RW, et. al. (1990). The Incidence of Fall Injury Events Among the Elderly in a Defined Population. *American Journal of Epidemiology* 131,1028-1037.

Schaffner W, et. al. (1983). Improving Antibiotic Prescribing in Office Practice: A Controlled Trial of Three Educational Methods. JAMA 250(13), 1728-1732.

Schauberger CW, et. al. (1996). A Quality Improvement Project to Increase the Use of Postmenopausal Hormonal Replacement Therapy. Wisconsin Medical Journal October , 697-701.

Schectman JM, et. al. (1995). The Effect of an Education and Feedback Intervention on Group-Model and Network-Model Health Maintenance Organization Physician Prescribing Behavior. Medical Care 33(2), 139-144.

Schoenberger JA. (1991). Drug-Induced Orostatic Hypotension. *Drug Safety* 6, 402-407.

Schroeder SA, et. al. (1984). The Failure of Physician Education as a Cost Containment Strategy. JAMA 252(2), 225-230.

Schwartz RK, Soumerai SB, and Avorn J. Physician Motivations for Nonscientific Drug Prescribing. Social Science and Medicine 28(6), 577-582.

Sever P. (1989). Combination Therapy with Calcium-Entry Blockers and Beta-Adrenoreceptor Atangonists in Hypertension. *Cardiovascular Drugs and Therapy* 3, 327-332.

Shorr RI and Robin DW. (1994). Rational Use of Benzodiazepines in the Elderly. *Drugs Aging* 4, 9-20.

Silvestry FE, and Kimmel SE. (1996). Calicum-channel Blockers in Ischemic Heart Disease. *Curr Opin Cardiol* 11, 434-439.

Simon G.E., VonKorff M., Heiligenstein J.H., Revicki D.A., Grothaus L., Katon W. and Wagner E.H. (1996). *JAMA* 275, 1897-1902.

Simon G.E., VonKorff M., Wagner, E.H. and Barlow W. (1993). Patterns of Antidepressant Use in Community Practice. *Gen Hosp Psychiatry* 15, 399-408.

Small RE, and Gonzalez ER. (1994). Managing the Elderly Patient with Peptic Ulcer Disease. *Hospital Formulary* 29 (suppl 1), S4-S11.

Sobel KG, McCart GM. (1983). Drug Use and Accidental Falls in an Intermediate Care Facility. *Drug Intell Clin Pharm* 17, 539-542.

The SOLVD Investigators. (1992). Effect of Enalapril on Mortality and Developement of Heart Failure in Asymptomatic Patients with Reduced Left Ventricular Ejection Fractions. *New England Journal of Medicine* 327, 685-691.

Sorock GS, and Shimkin EE. (1988). Benzodiazepine Sedatives and the Risk of Falling in a Community-Dwelling Elderly Cohort. *Archives of Internal Medicine* 148, 2441-2444.

Soumerai SB, and Avorn J. (1986) Economic and Policy Analysis of University-Based Drug "Detailing". Medical Care 24(4), 313-331.

Soumerai SB, and Avorn J. (1990). Principles of Educational Outreach ("Academic Detailing") to Improve Clinical Decision Making. JAMA 263(4), 549-556.

Soumerai SB, and Avorn J. (1984). Efficacy and Cost-Containment in Hospital Pharmacotherapy: State of the Art and Future Directions. Milbank Memorial Fund Quarterly 62(3), 447-474.

Soumerai SB, and Lipton HL. (1995). Computer-Based Drug Utilization Review --Risk, Benefit, or Boondoggle? The New England Journal of Medicine 332(24), 1641-1645.

Soumerai SB, McLaughlin TJ, and Avorn J. (1989). Improving Drug Prescribing in Primary Care: A Critical Analysis of the Experimental Literature. The Milbank Quarterly 67(2), 268-317.

Spangler DL. (1989). Review of Side Effects Associated with Beta Agonists. *Ann of Allergy* 62, 59-62.

Spore DL, et. al. (1997). Inappropriate Drug Prescriptions for Elderly Residents of Board and Care Facilities. American Journal of Public Health 87(3), 404-409.

Steinweg KK. (1997). The Changing Approach to Falls in the Elderly. *American Family Physician* 56, 1815-1823.

Stergachis A, et. al. (1987). Effect of Clinical Pharmacists on Drug Prescribing in a Primary-Care Clinic. American Journal of Hospital Pharmacy 44, 525-528.

Sulick JA, and Pathak DS. (1996). The Perceived Influence of Clinical Pharmacy Services on Physician Prescribing Behavior: A Matched-Pair Comparison of Pharmacists and Physicians. Pharmacotherapy 16(6), 1133-1141.

Thayssen P., Bjerre M., Dragh-Sorensen P., Moller M., Petersen O.L., et al. (1981). Cardiovascular Effects of Imipramine and Nortiptyline in Elderly Patients. *Psychopharmacology* 74, 360-64.

Tinetti ME, and Williams CS. (1997). Falls, Injuries Due to Falls, and the Risk of Admission to a Nursing Home. *New England Journal of Medicine* 337,1279-1284.

Tollefson, G.D. (1993). Adverse Drug Reactions/Interactions in Maintenance Therapy. *J Clicical Psychiatry* 54, 48-58.

Triggle DJ. (1995). Biochemical and Pharmacologic Differences among Calcium Channel Antagonists: Clinical Implications. In Calcium Antagonists in Clinical Medicine. (Ed.) Epstein M. Philadelphia: Hanley and Belfus.

Urton MM. (1991). A Community Home Inspection Approach to Preventing Falls Among the Elderly. *Public Health Reports* 106, 192-195.

Vasey FB. (1994). Using a Prodrug Approach to Reduce GI Toxicity in the Treatment of Arthritis. *Hospital Formulary* 29,289-98.

Vedsted P, Nielsen JN, and Olesen F. (1997). Does a computerized price comparison module reduce prescribing costs in general practice? Family Practice 14(3), 199-203.

Veith, R.C., Murray, A.R., Caldwell, J.H., Barnes, R.F., Gumbrecht, G., and Ritchie, J.L. (1982). Cardiovascular Effects of Tricyclic Antidepressants in Depressed Patients with Chronic Heart Disease. *New England Journal of Medicine* 306, 954-959.

Weedle PB, Poston JW, and Parich PA. (1988). The Use of Digoxin in 55 Residential Homes for Elderly People. *Postgraduate Medical Journal* 64, 292-296.

Weiner DK, Hanlon JT, and Studenski SA. (1998). Effects of Central Nervous System Polypharmacy on Falls Liability in Community-Dwelling Elderly. *Gerontology* 44(4), 217-221.

White H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica* 48, 149-170.

Wickham C, et. al. (1989). Muscle Strength, Activity, Housing and the Risk of Falls in Elderly People. *Age Aging* 18, 47-51.

Willett LR, Carson JL, and Strom BL. (1994). Epidemiology of Gastrointestinal Damage Associated with Nonsteroidal Anti-inflammatory Drugs. Division of General Internal Medicine, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical S. Drug Saf 10(2), 170-81.

Winniford MD, et. al. (1985). Propranolol-Verapamil Versus Propranolol-Nifedipine in Severe Angina Pectoris of Effort: A Randomized, Double-blind Crossover Study. *American Journal of Cardiology 55*, 281-285.

Wofford JL, et. al. (1992). Lack of Age-Related Differences in the Clinical Presentation of Digoxin Toxicity. *Archives of Internal Medicine* 152, 2261-2264.

Yates YM, Rupp MT, and Schondelmeyer SW. (1991). A Drug Utilization Evaluation Primer: Conceptual and Operational Aspects. In: The Contemporary HMO: Managing Care in the Modern Market. Proceedings of the 41st Annual Group Health Institute. New York, N.Y.

Yusuf S. (1995). Calcium Antagonists in Coronary Artery Disease and Hypertension. Time for Reevaluation? *Circulation* 92, 1079-1082.

Yusuf S, Held P, and Furberg C. (1991). Update of Effects of Calcium Antagonists in Myocardial Infarction or Angina in Light of the Second Danish Verapamil Infarction Trial (DAVIT-II) and other Recent Studies. *American Journal of Cardiology* 67, 1295-1297.

Zimmer, R., Gieschke, R., Fischbach, R., and Gasic.S. (1990). Interaction Studies with Moclobemide. *Acta Psychiatrica Scandinavica, Supplementum* 360, 84-86.

Zimmerman DR, et. al. (1994). Evaluation of a DUR Intervention: A Case Study of Histamine Antagonists. *Inquiry* 31: 89-101.

CMS LIBRARY
3 8095 00005838 4